# NAVAL POSTGRADUATE SCHOOL
## Monterey, California




# THESIS

## MATRIX ALGEBRA

by

Anthony R. Delatorre and William K. Cooke

June 1998

Advisor: Carlos F. Borges
Co-Advisor: Craig W. Rasmussen

**Approved for public release; Distribution is unlimited.**

# REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188



| 1. AGENCY USE ONLY (*Leave blank*) | 2. REPORT DATE June 1998 | 3. REPORT TYPE AND DATES COVERED Master's Thesis |
|---|---|---|

| | |
|---|---|
| 4. TITLE AND SUBTITLE MATRIX ALGEBRA | 5. FUNDING NUMBERS |
| 6. AUTHORS Delatorre, Anthony R. Cooke, William K. | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey CA 93943-5000 | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER |

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

| | |
|---|---|
| 12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited. | 12b. DISTRIBUTION CODE |

13. ABSTRACT(*maximum 200 words*)

This thesis is designed to act as an instructor's supplement for refresher matrix algebra courses at the Naval Postgraduate School (NPS). The need for a beginning matrix algebra supplement is driven by the unique circumstances of most NPS students. Most military students attend NPS several years after receiving their undergraduate degrees. This supplement, unlike most college textbooks, bridges the gap between the student's educational lay-off and the rigors of mathematically oriented degrees such as applied math, operations research and engineering. By reviewing the fundamental concepts of vectors and matrices, and performing basic operations with them, the student quickly develops the background needed in NPS's demanding curriculums. This supplement focuses on matrix and vector operations, linear transformations, systems of linear equations, and computational techniques for solving systems of linear equations. The goal is to enhance current matrix algebra textbooks and help the beginning student build a foundation for higher level engineering and mathematics based courses.

| | |
|---|---|
| 14. SUBJECT TERMS Introduction to Matrix Algebra. | 15. NUMBER OF PAGES 147 |
| | 16. PRICE CODE |

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | UL |

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

Approved for public release; distribution is unlimited

# MATRIX ALGEBRA


Anthony R. Delatorre
Lieutenant, United States Navy
B.S., University of Arizona, 1991

William K. Cooke
Lieutenant, United States Navy
B.E.E., Georgia Institute of Technology, 1989


Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN APPLIED MATHEMATICS**

from the

**NAVAL POSTGRADUATE SCHOOL**
**June 1998**

Authors: Anthony R. Delatorre

William K. Cooke

Approved by: Carlos F. Borges, Thesis Advisor

Craig W. Rasmussen, Co-Advisor

Walter M. Woods, Chair, Department of Mathematics

# ABSTRACT

This thesis is designed to act as an instructor's supplement for refresher matrix algebra courses at the Naval Postgraduate School (NPS). The need for a beginning matrix algebra supplement is driven by the unique circumstances of most NPS students. Most military students attend NPS several years after receiving their undergraduate degrees. This supplement, unlike most college textbooks, bridges the gap between the student's educational lay-off and the rigors of mathematically oriented degrees such as applied math, operations research and engineering. By reviewing the fundamental concepts of vectors and matrices, and performing basic operations with them, the student quickly develops the background needed in NPS's demanding curriculums. This supplement focuses on matrix and vector operations, linear transformations, systems of linear equations, and computational techniques for solving systems of linear equations. The goal is to enhance current matrix algebra textbooks and help the beginning student build a foundation for higher level engineering and mathematics based courses.

# TABLE OF CONTENTS

# INTRODUCTION TO MATRIX ALGEBRA

Matrix algebra is the study of algebraic operations on matrices and of their applications, primarily for solving systems of linear equations. Systems of linear equations appear often in fields that utilize mathematics to solve real-world problems. Although most of these real-world problems are so large that they would prove to be too difficult and time consuming to solve by hand, it is extremely important to understand the concepts underlying their solution. Once the techniques for solving small systems are understood, the methods can be applied to writing computer algorithms used to solve much larger systems. The purpose of this supplement therefore, is to develop the student's fundamental skills of matrix algebra problem solving at a very basic level. Where most matrix or linear algebra books concentrate on why, this supplement concentrates on how. That is, the computational aspect of matrix algebra is stressed instead of the derivation of the concept. Besides algebraic operations on matrices, we also discuss linear transformations. The focus of the supplement however, is to develop the skills necessary to solve systems of linear equations using several different methods. Whenever possible, examples using both real and complex numbers are used to get the student ready for the types of problems he or she is likely to encounter in follow on classes. Each chapter concludes with exercises designed to develop the students skills in the material just covered. The solutions can be found in Chapter Five. Students desiring a deeper understanding of matrix algebra will benefit greatly from the references used in writing this supplement.

Chapter One of the supplement deals with matrix notation and operations. The student learns how the elements of a matrix are arranged and the difference between square and rectangular matrices. We then introduce the basic algebraic operations for matrices: transpose, conjugate transpose, addition, subtraction, scalar multiplication, and matrix multiplication.

Chapter Two of the supplement is about vector operations and their geomet-

ric representations. The student learns that a vector can be regarded as a matrix consisting of either one row or one column. Therefore, the matrix operations of transpose, addition, and subtraction apply to vectors as well. Then we move on to vector multiplication. Since a vector can be thought of as a matrix, and we know how to multiply matrices, we can surely multiply a matrix times a vector. This introduces the matrix equation. The student then learns that there are two cases of vector multiplication-column $\times$ row and row $\times$ column. The dot product is introduced as a means of formally defining the inner-product. We then use the inner-product to generalize terms associated in two or three dimensions, such as length of a vector and the angle between two vectors. Using the inner-product, the student also learns how to project one vector onto another and determine if two vectors are orthogonal. The use of the inner-product allows us to deal with both real and complex vectors. We close the chapter by showing how a vector can be written as a linear combination of other vectors.

In Chapter Three, the student learns about linear transformations and how certain structured matrices like the identity matrix or permutation matrix can transform one vector into another vector. We then discuss matrix norms. Next, we move on to range and null space of a matrix. We finish up the chapter discussing a special type of matrix known as the elementary matrix. Elementary matrices are used to develop methods for solving systems of linear equations in the next chapter.

Chapter Four covers systems of linear equations and their solution. Here we define what is meant by a linear equation and by a system of linear equations and demonstrate several techniques used to solve them. The first three techniques are the method of substitution, Gaussian elimination, and Gauss-Jordan elimination. We also show the student how to find the inverse of a matrix. Because of their special form, triangular matrices are then studied as a lead in to our last solution method, the LU decomposition. We show two methods of finding the LU decomposition. We finish the chapter by looking at pivoting as it applies to computational efficiency.

This is a very important concept when systems are solved by computers.

This supplement is a compilation of several different works. Lecture notes from both advisors, a similar matrix algebra thesis, basic linear and matrix algebra books, and the authors first hand experiences of what topics at NPS seemed to need a bit more refresher attention.

The study of matrix algebra goes far beyond what has been presented in this supplement. The manual is designed to give students from varying backgrounds the necessary tools of matrix algebra needed to succeed in follow-on courses being taught at NPS.

# LIST OF REFERENCES

[1] Anton, Howard, *Elementary Linear Algebra*, John Wiley and Sons, New York, New York, 1994.

[2] Borges, Carlos F., *Lecture Notes on Linear Algebra*, Mathematics Department, Naval Postgraduate School, Monterey, California, 1995.

[3] Golub, Gene H. and Van Loan, Charles F., *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland, 1996.

[4] Rife, Susan A., *Matrix Algebra* (Thesis), Naval Postgraduate School, Monterey, California, 1996.

[5] Strang, Gilbert, *Linear Algebra and its Applications*, Harcourt Brace and Company, Orlando, Florida, 1988.

# INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center .................................... 2
   8725 John J. Kingman Road., Ste 0944
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library .................................................... 2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, CA 93943-5101

3. Chairman, Code MA ...................................................... 1
   Department of Mathematics
   Naval Postgraduate School
   1411 Cunningham Rd, Rm 341
   Monterey, CA 93943-5216

4. Professor Carlos Borges, Code MA/BC .................................... 2
   Department of Mathematics
   Naval Postgraduate School
   1411 Cunningham Rd, Rm 349
   Monterey, CA 93943-5216

5. Professor Craig Rasmussen, Code MA/RA .................................. 2
   Department of Mathematics
   Naval Postgraduate School
   1411 Cunningham Rd, Rm 348
   Monterey, CA 93943-5216

6. Lieutenant Anthony R. Delatorre ........................................ 1
   229 Worden St
   Portsmouth, RI 02871

7. Lieutenant William K. Cooke ............................................ 1
   445 Warley St
   Portsmouth, RI 02871

# PREFACE

The writing of this supplement is driven by the need for a manual to accompany current introductory matrix algebra textbooks. It is designed to provide additional examples and to clarify difficult concepts which are presented in the text. Students using this supplement are assumed to have varied educational backgrounds with different levels of knowledge, although it expected that all students have a basic facility with algebra.

This supplement is organized using the following format: We first discuss a concept, then formally define the concept, and then provide examples showing its use. Whenever applicable, examples using both real and complex numbers are used. At the end of each chapter there are exercises to reinforce the ideas that were covered in the chapter. The solutions to these exercises can be found in Chapter 5.

Chapter 1 introduces matrix notation and basic matrix operations. These operations are performed using both real and complex matrices. In Chapter 2 we talk about vector notation and operations, linear combinations, linear independence, and spanning sets. Chapter 3 is on linear transformations, including several specific examples such as the elementary matrices. Chapter 4 focuses on methods for solving systems of linear equations, these include substitution, Gaussian and Gauss-Jordan elimination, matrix inversion, and the LU decomposition.

# TABLE OF CONTENTS

# LIST OF FIGURES

# I. MATRIX OPERATIONS

We start by thinking of a *matrix* as a rectangular table of elements. A matrix composed of $n$ rows and $m$ columns is an element of $\mathcal{R}^{n\times m}$ if all of its elements are real numbers, and is an element of $\mathcal{C}^{n\times m}$ if any of its elements are complex numbers. The complex set is one that we will deal with quite frequently, since $\mathcal{R}^{n\times m} \subset \mathcal{C}^{n\times m}$. Upper case, Roman letters will be used to denote matrices. It is often useful to refer to a particular element in a matrix by its row and column position; the element in the $i$th row and $j$th column of a matrix A will be denoted by $a_{ij}$ or $\mathrm{A}_{ij}$.

**Example I.1** $\mathrm{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & a_{1m} \\ a_{2,1} & & & \cdot \\ \cdot & & & \cdot \\ a_{n1} & \cdot & \cdot & a_{nm} \end{bmatrix} \in \mathcal{C}^{n\times m}$ is a matrix with $n$ rows and $m$ columns.

**Example I.2** Given $\mathrm{A} = \begin{bmatrix} 1 & -i & -4 \\ 8 & 2i & 0 \\ 2 & 2-3i & 4 \end{bmatrix} \in \mathcal{C}^{3\times3}$, $a_{3,2} = 2-3i$ and is the element occupying the 3rd row and 2nd column position.

A *rectangular matrix* that has an equal number of rows and columns, like the matrix A in the previous example, is called a *square matrix.*

## A. MATRIX TRANSPOSE

One of the most useful operations that is performed on matrices is transposition. This operation turns the rows into columns and the columns into rows.

**Definition I.1** The *transpose* of a matrix A, denoted by $\mathrm{A}^T$, is defined by $(\mathrm{A}^T)_{ij} = a_{ji}$; that is, the $ij$th element of $\mathrm{A}^T$ is the $ji$th element of A.

**Example I.3** Given $\mathrm{A} = \begin{bmatrix} 1 & 4 & -6 & 0 \\ 3 & 8 & 12 & 3 \end{bmatrix}$, compute $\mathrm{A}^T$.

**Solution:**

$$A^T = \begin{bmatrix} 1 & 3 \\ 4 & 8 \\ -6 & 12 \\ 0 & 3 \end{bmatrix}.$$

An operation related to transposition is the construction of the conjugate transpose, or Hermitian transpose. The conjugate transpose of a matrix is constructed by transposing the matrix and taking the complex conjugate of all the elements, or, equivalently, by taking the complex conjugate and then transposing.

**Definition I.2** The *Hermitian transpose* or *conjugate transpose* of a matrix A, denoted by $A^H$, is defined by $(A^H)_{ij} = \overline{a}_{ji}$; or equivalently $A^H = \overline{A}^T = \overline{A^T}$.

**Example I.4** Given $A = \begin{bmatrix} 1 & 3i & 6 \\ -i & 2 & 4 \end{bmatrix}$, compute $A^T$ and $A^H$.

**Solution:**

$$A^T = \begin{bmatrix} 1 & -i \\ 3i & 2 \\ 6 & 4 \end{bmatrix} \text{ and } A^H = \begin{bmatrix} 1 & i \\ -3i & 2 \\ 6 & 4 \end{bmatrix}.$$

As this example demonstrates, the (Hermitian) transpose of an $n \times m$ matrix is an $m \times n$ matrix. Note that taking the (Hermitian) transpose of a matrix twice returns the original matrix. That is, $(A^T)^T = A$ and $(A^H)^H = A$.

**Example I.5** Given $A^H = \begin{bmatrix} 1 & i \\ -3i & 2 \\ 6 & 4 \end{bmatrix}$, compute $(A^H)^H$.

**Solution:**

$$(A^H)^H = \begin{bmatrix} 1 & 3i & 6 \\ -i & 2 & 4 \end{bmatrix},$$

which is our original A from the previous example.

**Example I.6** Given $A = \begin{bmatrix} 1 & 0 & 6 \\ 0 & 2 & 4 \end{bmatrix}$, compute $A^T$ and $A^H$.

**Solution:**

$$A^T = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 6 & 4 \end{bmatrix} \text{ and } A^H = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 6 & 4 \end{bmatrix}.$$

In the real case, the transpose and the Hermitian transpose are identical.

## B. MATRIX ADDITION, SUBTRACTION AND SCALAR MULTIPLICATION

Matrices become useful when simple, algebraic operations are utilized. When matrices are the same size, it is natural to add and subtract them. This is accomplished by performing the operation element-by-element on corresponding pairs of the matrix entries.

**Definition I.3** Given $A,B \in \mathcal{C}^{n\times m}$, the *sum,* $A \pm B$, is computed as

$$A \pm B = \begin{bmatrix} a_{1,1} \pm b_{1,1} & \cdot & \cdot & a_{1m} \pm b_{1m} \\ a_{1,2} \pm b_{1,2} & \cdot & \cdot & a_{2m} \pm b_{2m} \\ \cdot & & & \cdot \\ a_{n1} \pm b_{n1} & \cdot & \cdot & a_{nm} \pm b_{nm} \end{bmatrix}.$$

From this definition, we see that the resulting matrix is an $n \times m$ matrix, like the original matrices, and is an element of $\mathcal{C}^{n\times m}$. Because the reals are a subset of the complex numbers, the definition for addition and subtraction also applies to real matrices.

**Example I.7** Given $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} i & 12 \\ -7 & 2i \end{bmatrix}$, compute $A + B$ and $A - B$.

**Solution:**

$$\mathrm{A}+\mathrm{B} = \begin{bmatrix} 1+i & 2+12 \\ 3-7 & 4+2i \end{bmatrix} = \begin{bmatrix} 1+i & 14 \\ -4 & 4+2i \end{bmatrix}, \text{ and}$$

$$\mathrm{A}-\mathrm{B} = \begin{bmatrix} 1-i & 2-12 \\ 3-(-7) & 4-2i \end{bmatrix} = \begin{bmatrix} 1-i & -10 \\ 10 & 4-2i \end{bmatrix}.$$

If two matrices are not of the same dimensions, i.e., both from $\mathcal{R}^{n\times m}$ or $\mathcal{C}^{n\times m}$, then these operations are not defined and we say the matrices are *not compatibly sized* for addition and subtraction.

**Example I.8** Given $\mathrm{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\mathrm{B} = \begin{bmatrix} 1 & 0 \\ -3 & 2 \\ 6 & 4 \end{bmatrix}$, compute $\mathrm{A}+\mathrm{B}$.

**Solution:** A+B is not defined because A and B are not compatibly sized for addition.

Another useful operation is scalar multiplication, which is defined as follows.

**Definition I.4** Given a matrix $\mathrm{A} \in \mathcal{C}^{n\times m}$ and a scalar $\alpha \in \mathcal{C}$. The *scalar product*, $\alpha\mathrm{A}$, is formed by multiplying each element of A by $\alpha$:

$$\alpha\mathrm{A} = \begin{bmatrix} \alpha a_{1,1} & \alpha a_{1,2} & \cdot & \alpha a_{1m} \\ \alpha a_{2,1} & & & \cdot \\ \cdot & & & \cdot \\ \alpha a_{n1} & \cdot & \cdot & \alpha a_{nm} \end{bmatrix}.$$

**Example I.9** Given $\alpha = 3$ and $\mathrm{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4i \end{bmatrix}$, compute $\alpha\mathrm{A}$.

**Solution:**

$$\alpha\mathrm{A} = 3\begin{bmatrix} 1 & 2 \\ 3 & 4i \end{bmatrix} = \begin{bmatrix} 3\cdot 1 & 3\cdot 2 \\ 3\cdot 3 & 3\cdot 4i \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 9 & 12i \end{bmatrix}.$$

**Example I.10** Given $A = \begin{bmatrix} 1 & i \\ -3i & 2 \\ 6 & 4 \end{bmatrix}$ and $\alpha = -2 + i$, compute $\alpha A$.

**Solution:**

$$\alpha A = \begin{bmatrix} -2+i & -1-2i \\ 3+6i & -4+2i \\ -12+6i & -8+4i \end{bmatrix}.$$

When dealing with real numbers it is understood that $a + 0 = a$. Although we may not recall the exact definition, 0 is the additive identity. Yet what does $A + 0$ equal? Our definition for matrix addition dictates that we can only add matrices that are compatibly sized. If we make 0 a matrix filled with all zeros and ensure it is the same size as A, we see, $A + 0 = A$.

**Example I.11** Given $A = \begin{bmatrix} -2 & -3i \\ -8i & 3+4i \\ 8-9i & 0 \end{bmatrix}$, compute $A + 0$.

**Solution:**

$$A + 0 = \begin{bmatrix} -2 & -3i \\ -8i & 3+4i \\ 8-9i & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -2 & -3i \\ -8i & 3+4i \\ 8-9i & 0 \end{bmatrix}.$$

**Example I.12** Given $A = \begin{bmatrix} 8-9i & 5 & -1 & -3i \\ -2 & 0 & -8i & 7 \\ 3+4i & i & 6 & 0 \end{bmatrix}$, compute $A + 0$.

**Solution:**

$$A + 0 = \begin{bmatrix} 8-9i & 5 & -1 & -3i \\ -2 & 0 & -8i & 7 \\ 3+4i & i & 6 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 8-9i & 5 & -1 & -3i \\ -2 & 0 & -8i & 7 \\ 3+4i & i & 6 & 0 \end{bmatrix}.$$

**Example I.13** Given $A = \begin{bmatrix} 8-9i & 5 & -1 & -3i \\ -2 & 0 & -8i & 7 \\ 3+4i & i & 6 & 0 \end{bmatrix}$, compute $A + 1$.

**Solution:** A + 1 is not defined, since A and 1 are not compatibly sized for addition.

With the definitions of matrix addition, subtraction, and scalar multiplication, we can now list some important properties of matrices.

**Theorem I.1** *Given matrices* A, B, C $\in \mathcal{C}^{n\times m}$, *and scalars* $\alpha, \beta \in \mathcal{C}$, *the following properties of matrix addition, subtraction and scalar multiplication hold:*

1. $\mathrm{A}+\mathrm{B}=\mathrm{B}+\mathrm{A}$
2. $\mathrm{A}+(\mathrm{B}+\mathrm{C})=(\mathrm{A}+\mathrm{B})+\mathrm{C}$
3. $\alpha(\mathrm{A}+\mathrm{B})=\alpha\mathrm{A}+\alpha\mathrm{B}$
4. $(\alpha+\beta)\mathrm{A}=\alpha\mathrm{A}+\beta\mathrm{A}$
5. $(\alpha\beta)\,\mathrm{A}=\beta\,(\alpha\mathrm{A})$
6. A+0= A
7. A+(−A)=0

**Example I.14** Given $\mathrm{A}=\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$, $\mathrm{B}=\begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{bmatrix}$, and $\alpha \in \mathcal{C}$, determine whether $\alpha\;(\mathrm{A}+\mathrm{B})=\alpha\;\mathrm{A}+\alpha\;\mathrm{B}$.

**Solution:**

$$\begin{aligned}
\alpha\,(\mathrm{A}+\mathrm{B}) &= \begin{bmatrix} \alpha\,(a_{1,1}+b_{1,1}) & \alpha\,(a_{1,2}+b_{1,2}) & \alpha\,(a_{1,3}+b_{1,3}) \\ \alpha\,(a_{2,1}+b_{2,1}) & \alpha\,(a_{2,2}+b_{2,2}) & \alpha\,(a_{2,3}+b_{2,3}) \\ \alpha\,(a_{3,1}+b_{3,1}) & \alpha\,(a_{3,2}+b_{3,2}) & \alpha\,(a_{3,3}+b_{3,3}) \end{bmatrix} \\
&= \begin{bmatrix} \alpha a_{1,1}+\alpha b_{1,1} & \alpha a_{1,2}+\alpha b_{1,2} & \alpha a_{1,3}+\alpha b_{1,3} \\ \alpha a_{2,1}+\alpha b_{2,1} & \alpha a_{2,2}+\alpha b_{2,2} & \alpha a_{2,3}+\alpha b_{2,3} \\ \alpha a_{3,1}+\alpha b_{3,1} & \alpha a_{3,2}+\alpha b_{3,2} & \alpha a_{3,3}+\alpha b_{3,3} \end{bmatrix} \\
&= \begin{bmatrix} \alpha a_{1,1} & \alpha a_{1,2} & \alpha a_{1,3} \\ \alpha a_{2,1} & \alpha a_{2,2} & \alpha a_{2,3} \\ \alpha a_{3,1} & \alpha a_{3,2} & \alpha a_{3,3} \end{bmatrix} + \begin{bmatrix} \alpha b_{1,1} & \alpha b_{1,2} & \alpha b_{1,3} \\ \alpha b_{2,1} & \alpha b_{2,2} & \alpha b_{2,3} \\ \alpha b_{3,1} & \alpha b_{3,2} & \alpha b_{3,3} \end{bmatrix} = \alpha\mathrm{A}+\alpha\mathrm{B}.
\end{aligned}$$

This example is not a formal proof. However, we can use this technique to show that the given properties hold, and verifications of the remaining properties are left as exercises.

## C. MATRIX MULTIPLICATION

Matrix multiplication could be defined in the element-wise manner of addition and subtraction, but this leads to algebraic results of limited applicability. Instead, we define matrix multiplication in the following manner.

**Definition I.5** Given $A \in \mathcal{C}^{n \times p}$ and $B \in \mathcal{C}^{p \times m}$, the *product* $AB = C$ is defined as

$$(AB)_{ij} = \sum_{k=1}^{p} A_{ik}B_{kj} = C_{ij}.$$

**Example I.15** Given $A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \in \mathcal{C}^{2\times 2}$ and $B = \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \in \mathcal{C}^{2\times 2}$, compute AB.

**Solution:**

$$AB = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} = \begin{bmatrix} a_{1,1}b_{1,1} + a_{1,2}b_{2,1} & a_{1,1}b_{1,2} + a_{1,2}b_{2,2} \\ a_{2,1}b_{1,1} + a_{2,2}b_{2,1} & a_{2,1}b_{1,2} + a_{2,2}b_{2,2} \end{bmatrix} = C.$$

Here we can see that the $c_{1,1}$ entry is the sum of the products of the elements of the first row of A and the elements of the first column of B. The $c_{1,2}$ entry is the sum of the products of the elements of the 1st row of A and the elements of the second column of B. The next two entries of C, $c_{2,1}$ and $c_{2,2}$, are computed in the same manner. Therefore, we find the $ij$th element of C by finding the sum of the products of the elements $i$th row of A and the $j$th column of B.

**Example I.16** Given $A = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -2 & 0 \\ 1 & 6 & 4 \end{bmatrix}$, compute AB.

**Solution:**

$$\text{AB} = \begin{bmatrix} 1\cdot 1+3\cdot 1 & 1(-2)+3\cdot 6 & 1\cdot 0+3\cdot 4 \\ -1\cdot 1+2\cdot 1 & -1(-2)+2\cdot 6 & -1\cdot 0+2\cdot 4 \end{bmatrix} = \begin{bmatrix} 4 & 16 & 12 \\ 1 & 14 & 8 \end{bmatrix}.$$

**Example I.17** Given A $= \begin{bmatrix} -1-i & 3+2i \end{bmatrix}$ and B $= \begin{bmatrix} 1 & -2+i \\ -i & 6 \end{bmatrix}$, compute AB.

**Solution:**

$$\begin{aligned} \text{AB} &= \begin{bmatrix} (-1-i)1+(3+2i)(-i) & (-1-i)(-2+i)+(3+2i)6 \end{bmatrix} \\ &= \begin{bmatrix} 1-4i & 11+13i \end{bmatrix}. \end{aligned}$$

In the examples above, we found the product AB by following the simple rule of "finding the sum of products a row from matrix A and a column from matrix B". At this point, it should be fairly obvious that for the matrices to be compatible under this operation, the number of columns in A must equal the number of rows in B. These dimensions are called the *inner dimensions* of the matrix product. The two remaining dimensions are called the *outer dimensions* as illustrated below:

$$\overbrace{(3\times \underbrace{4)(4}_{inner} \times 2)}^{outer}.$$

**Example I.18** Given $A \in \mathcal{R}^{3\times 4}$ and $B \in \mathcal{R}^{4\times 2}$, determine whether A and B are compatibly sized for matrix multiplication, and if so, what are the dimensions of the resulting matrix?

**Solution:** The inner dimensions are equal, therefore A and B are compatibly sized for matrix multiplication. $AB \in \mathcal{R}^{3\times 2}$, since the outer dimensions are $3 \times 2$.

**Example I.19** Given $A \in \mathcal{R}^{2\times 2}$ and $B \in \mathcal{R}^{2\times 3}$, determine whether A and B are compatibly sized for matrix multiplication, and if so, what are the dimensions of the resulting matrix?

**Solution:** The inner dimensions are equal, therefore A and B are compatibly sized for matrix multiplication. $AB \in \mathcal{R}^{2\times 3}$, since the outer dimensions are $2 \times 3$.

**Example I.20** Given $A = \begin{bmatrix} 1 & i \\ -3i & 2 \\ 6 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & -i \\ 3 & 4 & i \end{bmatrix}$, compute AB.

**Solution:**

$$\begin{aligned} AB &= \begin{bmatrix} 1 & i \\ -3i & 2 \\ 6 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & -i \\ 3 & 4 & i \end{bmatrix} = \begin{bmatrix} 1+3i & 2+4i & -i-1 \\ -3i+6 & -6i+8 & -3+2i \\ 6+12 & 12+16 & -6i+4i \end{bmatrix} \\ &= \begin{bmatrix} 1+3i & 2+4i & -1-i \\ 6-3i & +8-6i & -3+2i \\ 18 & 28 & -2i \end{bmatrix}. \end{aligned}$$

The key factor that determines whether the operation of matrix multiplication is possible is whether or not the inner dimensions of A and B are equal, or in other words, that the two matrices are compatibly sized for forming the product AB. However, this does not necessarily mean that they are compatibly sized for forming the product BA. In the previous example, if we tried to find the matrix product BA, we could not since the inner dimensions of B and A are not equal. We say that BA is not defined. Even if the matrices are compatibly sized, AB does not necessarily equal BA.

**Example I.21** Given $A = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 5 \\ 9 & 0 \end{bmatrix}$, determine whether AB = BA.

**Solution:**

$$\begin{aligned} AB &= \begin{bmatrix} 1\cdot 4 + 3\cdot 9 & 1\cdot 5 + 3 \cdot 0 \\ -1\cdot 4 + 2\cdot 9 & -1\cdot 5 + 2\cdot 0 \end{bmatrix} = \begin{bmatrix} 31 & 5 \\ 14 & -5 \end{bmatrix}, \text{ and} \\ BA &= \begin{bmatrix} 4\cdot 1 + 5(-1) & 4\cdot 3 + 5\cdot 2 \\ 9\cdot 1 + 0(-1) & 9\cdot 3 + 0\cdot 2 \end{bmatrix} = \begin{bmatrix} -1 & 22 \\ 9 & 27 \end{bmatrix}. \end{aligned}$$

Clearly, these two products are not equal and, therefore, matrix multiplication as we have defined it *does not*, in general, commute.

**Example I.22** Given A $\in \mathcal{R}^{2\times 4}$ and B $\in \mathcal{R}^{4\times 2}$, AB $\in \mathcal{R}^{2\times 2}$ and BA $\in \mathcal{R}^{4\times 4}$. Here both AB and BA exist, but the resulting matrices do not have the same dimensions.

Some important properties of matrix operations follow.

**Theorem I.2** *Given matrices* A,B,C $\in \mathcal{C}^{n\times m}$, *and scalars* $\alpha, \beta \in \mathcal{C}$, *the following properties of matrix multiplication hold, provided that the indicated operations can be performed:*

1. $\mathrm{A(BC) = (AB)C}$
2. $\mathrm{A(B + C) = AB + AC}$
3. $\mathrm{(B + C)A = BA + CA}$
4. $\alpha(\mathrm{AB}) = (\alpha\mathrm{A})\mathrm{B} = (\mathrm{A}\alpha)\mathrm{B} = \mathrm{A}(\alpha\mathrm{B}) = \mathrm{A}(\mathrm{B}\alpha) = (\mathrm{AB})\alpha$
5. $(\alpha\beta)\mathrm{A} = \alpha(\beta\mathrm{A})$
6. $(\mathrm{A} + \mathrm{B})^H = \mathrm{A}^H + \mathrm{B}^H$
7. $(\mathrm{AB})^H = \mathrm{B}^H\mathrm{A}^H$

**Example I.23** Given $\mathrm{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$, $\mathrm{B} = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{bmatrix}$, and $\alpha, \beta \in \mathcal{C}$, determine whether $(\alpha\beta)\mathrm{A} = \alpha(\beta\mathrm{A})$.

**Solution:**

$$(\alpha\beta)\mathrm{A} = \begin{bmatrix} \alpha\beta a_{1,1} & \alpha\beta a_{1,2} & \alpha\beta a_{1,3} \\ \alpha\beta a_{2,1} & \alpha\beta a_{2,2} & \alpha\beta a_{2,3} \\ \alpha\beta a_{3,1} & \alpha\beta a_{3,2} & \alpha\beta a_{3,3} \end{bmatrix} = \alpha \begin{bmatrix} \beta a_{1,1} & \beta a_{1,2} & \beta a_{1,3} \\ \beta a_{2,1} & \beta a_{2,2} & \beta a_{2,3} \\ \beta a_{3,1} & \beta a_{3,2} & \beta a_{3,3} \end{bmatrix} = \alpha(\beta\mathrm{A}).$$

This example is not a formal proof. However, we can use this technique to verify that property I.2(5) holds and verifications of the remaining properties are left as exercises.

## D. EXERCISES

1. Given $A = \begin{bmatrix} 3 \\ -6 \\ 1 \end{bmatrix}$, $B = \begin{bmatrix} 1+3i & 2 \\ -6 & 0 \\ 1-i & -5i \end{bmatrix}$, and $C = \begin{bmatrix} 2-6i & 7-4i & 2 \\ -6+4i & -5+6i & -6+2i \\ 1 & 3-i & 0 \end{bmatrix}$,

write

a) $A^T$ b) $B^H$ c) $C^H$

2. Given $A = \begin{bmatrix} 5+i \\ 7 \\ -2i \end{bmatrix}$, $B = \begin{bmatrix} 1+i & -4 \\ 3 & -7i \end{bmatrix}$, and $C = \begin{bmatrix} 2 & 4i & 22 \\ -6 & -5 & -3+4i \\ 14+3i & -i & 10 \end{bmatrix}$,

write

a) $A^T$ b) $B^H$ c) $C^H$

3. Given $A = \begin{bmatrix} 3 \\ -6 \\ 1 \end{bmatrix}$, $B = \begin{bmatrix} 5 \\ -1 \\ 8 \end{bmatrix}$, $C = \begin{bmatrix} 4-2i & 3-4i \\ -2i & 0 \\ i & 5 \end{bmatrix}$, and $D = \begin{bmatrix} 7 & 0 \\ -8-4i & 2-3i \\ 1+5i & 9-6i \end{bmatrix}$,

compute

a) $A+B$ b) $B+D$ c) $D-C$ d) $C^H+D$

e) $2A-3B$ f) $iD^H+2C^T$

4. Given $A = \begin{bmatrix} 5 \\ 9 \\ 6 \end{bmatrix}$, $B = \begin{bmatrix} i \\ 6 \\ 10 \end{bmatrix}$, $C = \begin{bmatrix} 4-2i & 3-4i \\ -2i & 0 \\ i & 5 \end{bmatrix}$, and $D = \begin{bmatrix} i & -2+2i & 1 \\ -8 & 14i & 0 \end{bmatrix}$,

compute

a) $A-B$ b) $B+C$ c) $D^H+C$ d) $C^H-D$

e) $A+2B$ f) $D^H-iC^H$

5. Given matrices A $= \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$, B $= \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix}$, and C $= \begin{bmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{bmatrix}$, verify theorem I.1.

6. Given A $= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ and B $= \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -3 \\ 0 & 2 & 1 \end{bmatrix}$, show AB $\neq$ BA. Is this true in general?

7. Given A $= \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -3 \\ 0 & 2 & 1 \end{bmatrix}$, B $= \begin{bmatrix} 3 & 1 & 0 \\ 0 & 2 & 4 \\ 1 & 1 & -2 \end{bmatrix}$, C $= \begin{bmatrix} 2-6i & 7-4i & 2 \\ -6+4i & -5+6i & -6+2i \\ 1 & 3-i & 0 \end{bmatrix}$, and $a \in \mathcal{R}$, compute

a) BA    b) $\mathrm{B}^T\mathrm{C}$    c) $\mathrm{C}^H\mathrm{B}$    d) $a$(BA)    e) BA + BC

8. Given A $= \begin{bmatrix} 5 & 4 & 1 \\ 2 & 3 & 2 \end{bmatrix}$, B $= \begin{bmatrix} 2 & 4 & 6 \\ 1 & 2 & 3 \\ 1 & 5 & -1 \end{bmatrix}$, C $= \begin{bmatrix} 1 & 2+i \\ -3i & 3 \\ 1-2i & 0 \end{bmatrix}$, and $a \in \mathcal{R}$, compute

a) AB    b) $\mathrm{BA}^T$    c) $\mathrm{C}^H\mathrm{B}$    d) $a$(BC)    e) AB + C

9. Given A $= \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$, B $= \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix}$, and C $= \begin{bmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{bmatrix}$, verify theorem I.2.

# II. VECTOR OPERATIONS

We start by thinking of a *vector* as a special case of a "thin" matrix. Therefore, all of the rules for matrices will also hold true for vectors. A *column vector* is an element of $\mathcal{R}^{n\times 1}$. For simplicity of notation, we view column vectors as elements of $\mathcal{R}^n$ if all of the elements are real numbers, and of $\mathcal{C}^n$ if any of the elements are complex. The superscript $n$ denotes the number of elements in the column vector. Lower case, bold letters such as $\mathbf{v}$ and $\mathbf{w}$ will denote vectors. Regardless of whether $\mathbf{v}$ is a row vector or a column vector, the $i$th element of $\mathbf{v}$ will be denoted by $v_i$, where $i$ denotes the position of an element in the vector.

**Definition II.1** $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{bmatrix}$ is a *column vector,* and $\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdot & \cdot & w_n \end{bmatrix}$ is a *row vector.*

**Example II.1** Given $\mathbf{v} = \begin{bmatrix} 1 \\ 2i \\ 6 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 10-3i & -2+i & -3i & 9 \end{bmatrix}$, their third elements are $v_3 = 6$ and $w_3 = -3i$, respectively.

Transposition is simply the act of making a column vector a row vector or a row vector a column vector. The orders of the elements are unchanged.

**Definition II.2** The *transpose* of $\mathbf{v}$, denoted by $\mathbf{v}^T$, is defined by $(\mathbf{v}^T)_i = v_i$; that is, the $i$th element of the transpose of $\mathbf{v}$ is the conjugate of the $i$th element of $\mathbf{v}$.

**Example II.2** Given $\mathbf{w} = \begin{bmatrix} 12 & 0 & -3 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}$, compute $\mathbf{w}^T$ and $\mathbf{v}^T$.

**Solution:**

$$\mathbf{w}^T = \begin{bmatrix} 12 \\ 0 \\ -3 \end{bmatrix} \text{ and } \mathbf{v}^T = \begin{bmatrix} 1 & 2 & 6 \end{bmatrix}.$$

**Definition II.3** The *conjugate transpose* of $\mathbf{v}$, denoted by $\mathbf{v}^H$, is defined by $(\mathbf{v}^H)_i = \overline{v}_i$; that is, the $i$th element of the conjugate transpose of $\mathbf{v}$ is the conjugate of the $i$th element of $\mathbf{v}$.

**Example II.3** Given $\mathbf{v} = \begin{bmatrix} 1 \\ 2i \\ 6 \end{bmatrix}$, compute $\mathbf{v}^H$ and $\left(\mathbf{v}^H\right)^H$.

**Solution:**

$$\mathbf{v}^H = \begin{bmatrix} 1 & -2i & 6 \end{bmatrix} \text{ and } \left(\mathbf{v}^H\right)^H = \begin{bmatrix} 1 \\ 2i \\ 6 \end{bmatrix}.$$

Notice that transposing a vector twice returns the original vector. This means that $(\mathbf{v}^T)^T = \mathbf{v}$ and $(\mathbf{v}^H)^H = \mathbf{v}$. Unless otherwise noted, *all vectors should be interpreted to be column vectors.* A row vector will be written as the transpose of a column vector. In the next few examples, remember to think of vectors as skinny matrices.

**Example II.4** Given $\mathbf{v} = \begin{bmatrix} -2+4i & 1+3i \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} \frac{1}{2} & -6i \end{bmatrix}^T$, compute $\mathbf{v}+\mathbf{w}$ and $\mathbf{v}-\mathbf{w}$.

**Solution:**

$$\begin{aligned} \mathbf{v}+\mathbf{w} &= \begin{bmatrix} -2+1/2+4i \\ 1+i(3-6) \end{bmatrix} = \begin{bmatrix} -3/2+4i \\ 1-3i \end{bmatrix}, \text{ and} \\ \mathbf{v}-\mathbf{w} &= \begin{bmatrix} -2-1/2+4i \\ 1+i(3+6) \end{bmatrix} = \begin{bmatrix} -5/2+4i \\ 1+9i \end{bmatrix}. \end{aligned}$$

If two vectors are not from the same space, $\mathcal{R}^n$ or $\mathcal{C}^n$, then the operations of addition and subtraction are not defined and we say that the vectors are not compatibly sized for addition and subtraction.

**Example II.5** Given $\mathbf{v} = \begin{bmatrix} -2+4i & 1+3i & 4 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} \frac{1}{2} & -6i \end{bmatrix}^T$, compute $\mathbf{v}+\mathbf{w}$.

**Solution:** $\mathbf{v}+\mathbf{w}$ is not defined, since $\mathbf{v}$ and $\mathbf{w}$ are not compatibly sized for addition.

**Example II.6** Given $\mathbf{v} = \begin{bmatrix} 3 & -2 & 0 & 5-3i \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} -4 & 3i & -i & 1 \end{bmatrix}^T$, compute $\mathbf{v} - \mathbf{w}$.

**Solution:**

$$\mathbf{v} - \mathbf{w} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ 5-3i \end{bmatrix} - \begin{bmatrix} -4 \\ 3i \\ -i \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ -2-3i \\ i \\ 4-3i \end{bmatrix}.$$

**Example II.7** Given $\alpha = 2$ and $\mathbf{v} = \begin{bmatrix} 3 & 2i & -1 & 0 \end{bmatrix}^T$, compute $\alpha\mathbf{v}$.

**Solution:**

$$\alpha\mathbf{v} = 2 \begin{bmatrix} 3 \\ 2i \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 4i \\ -2 \\ 0 \end{bmatrix}.$$

**Example II.8** Given $\alpha = -2+i$ and $\mathbf{v} = \begin{bmatrix} -2i & 0 & -1+4i & -3-5i \end{bmatrix}^T$, compute $\alpha\mathbf{v}$.

**Solution:**

$$\alpha\mathbf{v} = (-2+i) \begin{bmatrix} -2i \\ 0 \\ -1+4i \\ -3-5i \end{bmatrix} = \begin{bmatrix} 2+4i \\ 0 \\ -2-5i \\ 11+7i \end{bmatrix}.$$

To further understand how these vector operations work, we might look at them geometrically. Most people have no trouble understanding vectors in the two-dimensional sense, as an arrow in the Cartesian plane. Just as we can add and subtract vectors algebraically, we can also "add" and "subtract" the arrows to create new arrows, or vectors. The following paragraphs will attempt to tie together the notion of vectors in n space and the geometric interpretation of the basic algebraic operations of addition, subtraction, and scalar multiplication.

In a simple two-dimensional case, a vector can be viewed as an arrow emanating from the origin. The elements of the vector consist of the coordinates of the head of the arrow, or, similarly, the distance moved horizontally and vertically from the origin. As shown in Figure 1, the vector **v** is represented by the coordinate pair (3,1), and **w** by the coordinate pair (1,3). We represent this pictorially as shown in Figure 1.

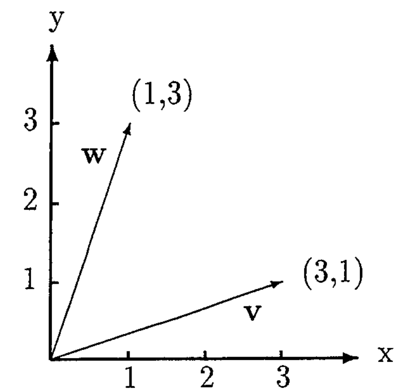


Figure 1. Two-Dimensional Vector Representation

$$\mathbf{v} = \begin{bmatrix} 3 & 1 \end{bmatrix}^T, \text{ and } \mathbf{w} = \begin{bmatrix} 1 & 3 \end{bmatrix}^T.$$

This is the two-dimensional case. We can then easily visualize the three-dimensional case. Given $\mathbf{v} = \begin{bmatrix} x & y & z \end{bmatrix}^T$, we represent **v** by an arrow from the origin to P$(x, y, z)$.

Vectors in two and three dimensions are easy to visualize, and we can translate geometric concepts such as length into vector operations. Although $n$ dimensions are difficult to visualize, we can still talk about addition, subtraction, scalar multiplication, and length of $n$-dimensional vectors. Notice again, that, as in Figure 1, if the order of a vector **v**'s elements are rearranged, a new vector **w** is formed that is entirely different from the original vector. So we see that the order of vector elements is important.

Now we look at vector addition in the Cartesian plane. To add any two vectors $\mathbf{v}$ and $\mathbf{w}$ geometrically, the vectors are first attached head to tail. In Figure 2, if we add $\mathbf{v}$ to $\mathbf{w}$, the resulting vector $\mathbf{v}+\mathbf{w}$ is computed by sliding a copy of vector $\mathbf{w}$ without changing its direction, over to the head of vector $\mathbf{v}$. Then draw a line from the origin to the head of the copied vector $\mathbf{w}$. This new line emanating from the origin is the solution vector $\mathbf{v}+\mathbf{w}$. If the two vectors are added analytically, the result will be the same as the geometric solution. The key to remember is that, when the copy of the second vector is moved, its head and tail must remain in the same relative positions as the original.

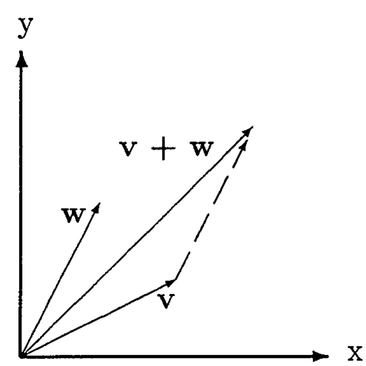


Figure 2. Vector Addition

Vector subtraction is performed in the same manner as addition. The only difference is that the negative of the second vector is used. Therefore, as shown in Figure 3, to compute $\mathbf{v}-\mathbf{w}$, we subtract vector $\mathbf{w}$ from $\mathbf{v}$ by taking the negative of vector $\mathbf{w}$, which reverses its direction, and then slide a copy of it over to the head of vector $\mathbf{v}$. This is the same procedure as adding two vectors. We are now just adding the negative of $\mathbf{w}$.

To understand scalar multiplication, it might help to look at Figure 4. Again, we have two vectors $\mathbf{v}$ and $\mathbf{w}$. Notice $\mathbf{v}+\mathbf{v}=2\mathbf{v}$. Therefore, scalar multiplication is just the act of stretching, shrinking or reversing the direction of the original vector.

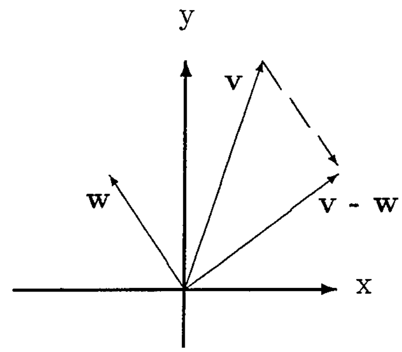


Figure 3. Vector Subtraction

If we multiply the vector **v** by the scalar 2, we see that it is stretched to twice its original length in the same direction. If we multiply vector **w** by the scalar $\frac{3}{2}$, we see that it is stretched so that the new vector is half again as long as the original **w**. If either of these vectors were multiplied by a negative scalar, the vector would reverse direction, and, depending on the magnitude of the scalar, the vector would be stretched or compressed.

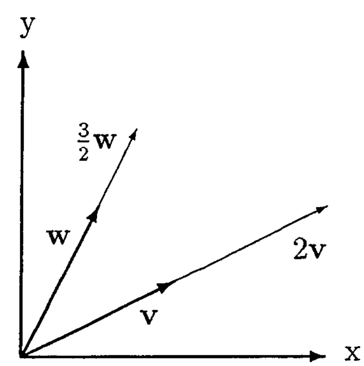


Figure 4. Scalar Multiplication of a Vector

With the definitions and geometric representation of vector addition, subtraction, and scalar multiplication, we can now show some important properties of

vectors.

**Theorem II.1** *Given the vectors* $\mathbf{v}, \mathbf{w}, \mathbf{u} \in \mathcal{C}^n$*, and the scalars* $\alpha, \beta \in \mathcal{C}$*, the following properties of vector addition, subtraction, and scalar multiplication hold:*

1. $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$
2. $\mathbf{v} + (\mathbf{w} + \mathbf{u}) = (\mathbf{v} + \mathbf{w}) + \mathbf{u}$
3. $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$
4. $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$
5. $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$
6. $\mathbf{v} + \mathbf{0} = \mathbf{v}$
7. $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
8. $1\mathbf{v} = \mathbf{v}$
9. $(-1)\mathbf{v} = -\mathbf{v}$
10. $0\mathbf{v} = \mathbf{0}$

**Example II.9** Given $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 \end{bmatrix}^T$, determine whether $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$.

**Solution:**

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \\ v_4 + w_4 \end{bmatrix} = \begin{bmatrix} w_1 + v_1 \\ w_2 + v_2 \\ w_3 + v_3 \\ w_4 + v_4 \end{bmatrix} = \mathbf{w} + \mathbf{v}.$$

We can easily extend the above example to any size vector. The proofs of the remaining properties are left as exercises.

## A. THREE SPECIAL CASES OF MATRIX MULTIPLICATION

Recall the definition for matrix multiplication: if $\mathrm{A} \in \mathcal{C}^{n \times p}$ and $\mathrm{B} \in \mathcal{C}^{p \times m}$, the product AB is defined as $(\mathrm{AB})_{ij} = \sum_{k=1}^{p} \mathrm{A}_{ik}\mathrm{B}_{kj}$. The first special case occurs when

we multiply a matrix and a column. The matrix multiplication definition becomes $(\mathrm{Ax})_i = \sum_{k=1}^{p} \mathrm{A}_{ik}\mathrm{x}_{k1}$. The equation $\mathrm{Ax} = \mathbf{b}$ is called the *matrix equation.*

**Example II.10** Given $\mathrm{A} = \begin{bmatrix} -2 & 0 \\ 1 & 4 \\ -3 & 5 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, compute $\mathbf{b} = \mathrm{Ax}$.

**Solution:**

$$\mathrm{Ax} = \begin{bmatrix} -2 & 0 \\ 1 & 4 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 2 \end{bmatrix} = \mathbf{b}.$$

The second special case occurs when we left-multiply a single column by a row. The matrix multiplication definition becomes $\mathbf{x}^T\mathbf{y} = \sum_{k=1}^{p} \mathrm{A}_{1k}\mathrm{B}_{k1} = a$. Notice that the matrices must still be compatibly sized. The result is a scalar, which leads to the name for this special case, i.e., the *scalar product* (or $a$).

**Example II.11** Given $\mathbf{x} = \begin{bmatrix} 3 & -4 & 0 & -2 \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} 4 & 1 & -5 & -3 \end{bmatrix}^T$, compute $\mathbf{x}^T\mathbf{y}$.

**Solution:** $\mathbf{x}^T$ and $\mathbf{y}$ are compatibly sized for matrix multiplication since $\mathbf{x} \in \mathcal{R}^{1\times 4}$ and $\mathbf{y} \in \mathcal{R}^{4\times 1}$. Therefore,

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} 3 & -4 & 0 & -2 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ -5 \\ -3 \end{bmatrix} = (12 - 4 + 0 + 6) = 14.$$

The third special case occurs when we left-multiply a row by a column. This result is called the *outer product.* The matrix multiplication definition becomes $\mathbf{x}\mathbf{y}^T = \sum_{k=1}^{p} \mathrm{A}_{i1}\mathrm{B}_{1i} = \mathrm{C}$. Again, notice that the matrices must still be compatibly sized. So we get a square matrix.

**Example II.12** Given $\mathbf{x} = \begin{bmatrix} 3 & -4 & 0 & -2 \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} 4 & 1 & -5 & -3 \end{bmatrix}^T$, compute $\mathbf{x}\mathbf{y}^T$.

**Solution:**

$$\mathbf{x}\mathbf{y}^T = \begin{bmatrix} 4 \\ 1 \\ -5 \\ -3 \end{bmatrix} \begin{bmatrix} 3 & -4 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 12 & -16 & 0 & -8 \\ 3 & -4 & 0 & -2 \\ -15 & 20 & 0 & 10 \\ -3 & 12 & 0 & 6 \end{bmatrix}.$$

In the last two examples we computed the scalar product $\mathbf{x}^T\mathbf{y}$ and the outer product $\mathbf{x}\mathbf{y}^T$, using the definition for matrix multiplication. The results were a scalar and a matrix, respectively. In the dot product section we will examine columns and rows more carefully.

## B. VECTOR NORM

As we have already discussed, we can think of a vector as a point in $n$-space where the coordinates of the point are given by the elements of the vector. For example, we can think of the vector $\mathbf{v} = \begin{bmatrix} 3 & 2 \end{bmatrix}^T$ as a point in the Cartesian plane with coordinates $(3, 2)$. It is also common to think of a vector as the directed line segment joining the origin and a coordinate point. Using the Pythagorean theorem, we can then find the length of the vector, or the distance from the origin to the point.

**Example II.13** Given $\mathbf{v} = \begin{bmatrix} 3 & 2 \end{bmatrix}^T$, compute the length of $\mathbf{v}$.

**Solution:** The length of $\mathbf{v}$ is $\sqrt{3^2 + 2^2} = \sqrt{13}$.

Notice that this looks similar to the special case of a row times a column,

$$\mathbf{v}^T\mathbf{v} = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = 3^2 + 2^2 = 13.$$

All we have to do is take the square root of this special case and we have the length.

**Definition II.4** Given a point (x,y) in the plane, the *distance* from the origin to the point is $\sqrt{x^2 + y^2}$.

Since $\mathbf{v}^T\mathbf{v}$ is the square of the length of $\mathbf{v}$, it seems natural to define the length of a vector using this special case.

**Definition II.5** Given a vector $\mathbf{v} \in \mathcal{C}^n$, the *Euclidean length,* or 2-*norm,* of $\mathbf{v}$, denoted by $\|\mathbf{v}\|_2$, is

$$\|\mathbf{v}\|_2 = \sqrt{\mathbf{v}^H\mathbf{v}} = \left( \begin{bmatrix} v_1 & v_2 & \cdot & \cdot & v_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{bmatrix} \right)^{\frac{1}{2}} = \sqrt{(v_1^2 + v_2^2 + \cdots + v_n^2)}.$$

**Example II.14** Given $\mathbf{v} = \begin{bmatrix} -3 & 1 & 4 \end{bmatrix}^T$, compute $\|\mathbf{v}\|_2$.

**Solution:**

$$\|\mathbf{v}\|_2 = \sqrt{\mathbf{v}^H\mathbf{v}} = \left( \begin{bmatrix} -3 & 1 & 4 \end{bmatrix} \begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix} \right)^{\frac{1}{2}} = \sqrt{9+1+16} = \sqrt{26}.$$

**Example II.15** Given $\mathbf{v} = \begin{bmatrix} a+bi \\ c+di \end{bmatrix}$ and $a, b, c, d \in \mathcal{R}$, compute $\mathbf{v}^H\mathbf{v}$.

**Solution:**

$$\begin{aligned} \mathbf{v}^H\mathbf{v} &= \begin{bmatrix} a-bi & c-di \end{bmatrix} \begin{bmatrix} a+bi \\ c+di \end{bmatrix} = (a-bi)(a+bi) + (c-di)(c+di) \\ &= a^2 - b^2i^2 + c^2 - d^2i^2 = a^2 + b^2 + c^2 + d^2. \end{aligned}$$

This operation always produces a real scalar, and therefore agrees with our notion of length.

**Example II.16** Given $\mathbf{v} = \begin{bmatrix} -2+4i & 1+3i & 5 & -3-i \end{bmatrix}^T$, compute $\|\mathbf{v}\|_2$.

**Solution:**

$$\|\mathbf{v}\|_2 = \sqrt{\mathbf{v}^H\mathbf{v}} = \left( \begin{bmatrix} -2-4i & 1-3i & 5 & -3+i \end{bmatrix} \begin{bmatrix} -2+4i \\ 1+3i \\ 5 \\ -3-i \end{bmatrix} \right)^{\frac{1}{2}}$$

$$= \sqrt{20+10+25+10} = \sqrt{65}.$$

It is important to understand there are numerous other norms which are all abstractions of the notion of length. One of these norms is the $p$-norm. The 2-norm is a specific case ($p = 2$) of a $p$-norm, which is defined as follows.

**Definition II.6** Given a vector $\mathbf{v} \in \mathcal{C}^n$, the *p-norm* of $\mathbf{v}$, denoted by $\|\mathbf{v}\|_p$, is

$$\|\mathbf{v}\|_p = \left( \sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}}, \text{ where } p \geq 1.$$

**Example II.17** Given $\mathbf{v} = \begin{bmatrix} -2 & 1 & 5 & -3 \end{bmatrix}^T$, compute $\|\mathbf{v}\|_1$.

**Solution:**

$$\|\mathbf{v}\|_1 = |-2| + |1| + |5| + |-3| = 11.$$

**Example II.18** Given $\mathbf{v} = \begin{bmatrix} -2 & 1 & 5 & -3 \end{bmatrix}^T$, compute $\|\mathbf{v}\|_2$.

**Solution:**

$$\|\mathbf{v}\|_2 = \left(|-2|^2 + |1|^2 + |5|^2 + |-3|^2\right)^{\frac{1}{2}} = \sqrt{39}.$$

**Example II.19** Given $\mathbf{v} = \begin{bmatrix} -2 & 1 & 5 & -3 \end{bmatrix}^T$, compute $\|\mathbf{v}\|_\infty$.

**Solution:**

$$\|\mathbf{v}\|_\infty = \max_{1 \leq i \leq n} | v_i | = \max\{|-2|, |1|, |5|, |-3|\} = 5.$$

We found $\|\mathbf{v}\|_\infty$ by letting $p \to \infty$ in $\left(\sum_{i=1}^{n} |v_i|^p\right)^{\frac{1}{p}}$, and taking the limit.

## C. THE DOT PRODUCT

Vector addition and subtraction in the plane seem natural, and we can visualize how a solution is found. On the other hand, we can define the notion of vector multiplication in many different ways. The *dot product,* also known as the *scalar product*, is one way to define vector multiplication.

**Definition II.7** Given two vectors $\mathbf{v}, \mathbf{w} \in \mathcal{R}^n$, the *dot product* of $\mathbf{v}$ with $\mathbf{w}$, denoted by $\mathbf{v} \cdot \mathbf{w}$, is

$$\mathbf{v} \cdot \mathbf{w} = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ w_n \end{bmatrix} = (v_1 w_1 + v_2 w_2 + \cdots v_n w_n).$$

**Example II.20** Given $\mathbf{v} = \begin{bmatrix} -4 & 1 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} 2 & -6 \end{bmatrix}^T$, compute $\mathbf{v} \cdot \mathbf{w}$.

**Solution:**

$$\mathbf{v} \cdot \mathbf{w} = (-4)(2) + (1)(-6) = -8 - 6 = -14.$$

Note that, when we take the dot product of two vectors, the result is a scalar, hence the name scalar product. The definition of dot product is valid for real vectors, but we will need a different definition in order to deal with complex vectors. The dot product is a specific case of an *inner product.* The next definition is also a special case of an inner product. We will use this definition by default when we are talking about the inner product.

**Definition II.8** Given two vectors $\mathbf{v}, \mathbf{w} \in \mathcal{R}^n$, the *inner product* of $\mathbf{v}$ with $\mathbf{w}$, denoted by $\mathbf{v}^T\mathbf{w}$, is

$$\mathbf{v}^T\mathbf{w} = \begin{bmatrix} v_1 & v_2 & \cdot & \cdot & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ w_n \end{bmatrix} = (v_1w_1 + v_2w_2 + \cdots + v_nw_n).$$

This is the dot product, again. But this definition is a natural extension of the definition of matrix multiplication. The inner product for complex vectors is similar to that of the reals, except that the Hermitian transpose of $\mathbf{v}$ is used instead of the transpose of $\mathbf{v}$, as shown in the following definition.

**Definition II.9** Given two vectors $\mathbf{v}, \mathbf{w} \in \mathcal{C}^n$, the *inner product* of $\mathbf{w}$ with $\mathbf{v}$, denoted by $\mathbf{w}^H\mathbf{v}$, is given by

$$\mathbf{w}^H\mathbf{v} = \begin{bmatrix} \overline{w_1} & \overline{w_2} & \cdot & \cdot & \overline{w_n} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{bmatrix} = (\overline{v_1}w_1 + \overline{v_2}w_2 + \cdots + \overline{v_n}w_n).$$

**Example II.21** Given $\mathbf{v} = \begin{bmatrix} -2+4i & 1+3i \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} 2 & -i \end{bmatrix}^T$, compute $\mathbf{w}^H\mathbf{v}$.

**Solution:**

$$\begin{aligned} \mathbf{w}^H\mathbf{v} &= \begin{bmatrix} 2 & i \end{bmatrix} \begin{bmatrix} -2+4i \\ 1+3i \end{bmatrix} = (2)(-2+4i) + (i)(1+3i) \\ &= -4 + 8i + i - 3 = -7 + 9i. \end{aligned}$$

What happens in this example if we reverse the order of the vectors? That is, if we compute $\mathbf{v}^H\mathbf{w}$? If the vectors are real, we get the same result. But notice what happens when the vectors are complex, as in the following example.

**Example II.22** Given $\mathbf{v} = \begin{bmatrix} -2+4i & 1+3i \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} 2 & -i \end{bmatrix}^T$, compute $\mathbf{v}^H\mathbf{w}$.

**Solution:**

$$\begin{aligned} \mathbf{v}^H\mathbf{w} &= \begin{bmatrix} -2-4i & 1-3i \end{bmatrix} \begin{bmatrix} 2 \\ -i \end{bmatrix} = (-2-4i)\cdot(2) + (1-3i)\cdot(-i) \\ &= -4-8i-i-3 = -7-9i. \end{aligned}$$

This is the complex conjugate of the product in the previous example, and demonstrates that, in general, $\mathbf{v}^H\mathbf{w} \neq \mathbf{w}^H\mathbf{v}$. However, $\mathbf{w}^H\mathbf{v} = \overline{\mathbf{v}^H\mathbf{w}}$ always holds. If one or both vectors are complex, then the two different orders of computing the inner product lead to results that are complex conjugates of each other. Notice that the definition for the inner product of complex vectors can be used for real vectors.

**Example II.23** Given $\mathbf{v} = \begin{bmatrix} -2+4i & 1+3i & 5 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} 2 & -i \end{bmatrix}^T$, compute $\mathbf{w}^T\mathbf{v}$.

**Solution:** $\mathbf{w}^T\mathbf{v}$ is not defined since the two vectors are not compatibly sized for vector multiplication.

**Example II.24** Given $\mathbf{v} = \begin{bmatrix} -2+4i & 1+3i & 5 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} -5i & -4-i & 6 \end{bmatrix}^T$, compute $\mathbf{w}^H\mathbf{v}$.

**Solution:**

$$\begin{aligned} \mathbf{w}^H\mathbf{v} &= \begin{bmatrix} 5i & -4+i & 6 \end{bmatrix} \begin{bmatrix} -2+4i \\ 1+3i \\ 5 \end{bmatrix} \\ &= 5i(-2+4i) + (-4+i)(1+3i) + (6)(5) = 3-21i. \end{aligned}$$

Now that we have defined the inner product, we can show some properties of vector products.

**Theorem II.2** *Given the vectors* $\mathbf{v}, \mathbf{w}, \mathbf{u} \in \mathcal{C}^n$*, and the scalars* $\alpha, \beta \in \mathcal{C}$*, the following properties of vector multiplication hold:*

1. $\mathbf{v}^T(\mathbf{w} + \mathbf{u}) = \mathbf{v}^T\mathbf{w} + \mathbf{v}^T\mathbf{u}$ (left distributive law)
2. $(\mathbf{w}^T + \mathbf{u}^T)\mathbf{v} = \mathbf{w}^T\mathbf{v} + \mathbf{u}^T\mathbf{v}$ (right distributive law)
3. $\alpha(\mathbf{v}^T\mathbf{w}) = (\alpha\mathbf{v}^T)\mathbf{w} = \mathbf{v}^T(\alpha\mathbf{w})$

The proofs are left as exercises.

## D. ORTHOGONALITY

Orthogonality is another term best explained in a specific and easily visualized case. In $\mathcal{R}^2$ and $\mathcal{R}^3$ we say that two vectors $\mathbf{v}$ and $\mathbf{w}$ are *orthogonal* if they are perpendicular, and write $\mathbf{v} \perp \mathbf{w}$. As the dimensions of $\mathbf{v}$ and $\mathbf{w}$ increase, it is difficult to visualize perpendicular vectors. If we talk about complex vectors it is nearly impossible to visualize perpendicular vectors. However, we can define "perpendicular" vectors of any dimension, real or complex. We will use the inner product to determine whether two vectors satisfy our definition.

**Definition II.10** Vectors $\mathbf{v}, \mathbf{w} \in \mathcal{C}^n$ are *orthogonal* if $\mathbf{v}^H\mathbf{w} = 0$.

**Example II.25** Given $\mathbf{v} = \begin{bmatrix} 2 & 4 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} 2 & -1 \end{bmatrix}^T$, determine whether $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

**Solution:**

$$\mathbf{v}^T\mathbf{w} = (2)(2) + (4)(-1) = 0.$$

Therefore, $\mathbf{v}$ is orthogonal to $\mathbf{w}$. This can be quickly verified by drawing the vectors in the Cartesian plane. If we look at Figure 5, we see that $\mathbf{v} = \begin{bmatrix} 2 & 4 \end{bmatrix}^T$ is orthogonal to $\mathbf{w} = \begin{bmatrix} 2 & -1 \end{bmatrix}^T$ since they are separated by 90 degrees. Yet, $\mathbf{u} = \begin{bmatrix} -1 & 3 \end{bmatrix}^T$ is not orthogonal to either $\mathbf{v}$ or $\mathbf{w}$. This can be verified by computing the inner product or

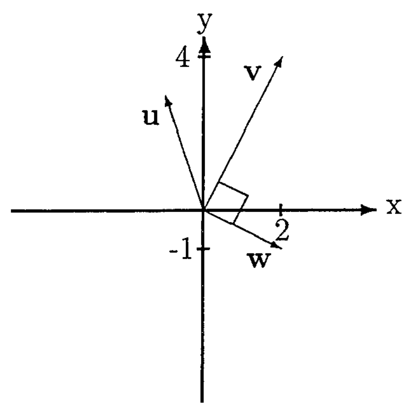


Figure 5. Vector Orthogonality

by looking at the vectors geometrically. Although it is difficult to visualize more than three dimensions, using the definition of orthogonality we can determine whether two vectors having the same number of elements are "perpendicular" or not.

**Example II.26** Given $\mathbf{v} = \begin{bmatrix} -2 & 1 & 5 & -3 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} -5 & -4 & 7 & 6 \end{bmatrix}^T$, determine whether $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

**Solution:** $\mathbf{v}^T\mathbf{w} = -2(-5) + 1(-4) + 5(7) - 3(6) = 23$. Therefore, $\mathbf{v}$ and $\mathbf{w}$ are not orthogonal.

**Example II.27** Given $\mathbf{v} = \begin{bmatrix} 4 & 0 & 2 & 4 & -7 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} -2 & 9 & 6 & -1 & 0 \end{bmatrix}^T$, determine whether $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

**Solution:** $\mathbf{v}^T\mathbf{w} = 4(-2) + 0(9) + 2(6) + 4(-1) - 7(0) = 0$. Therefore, $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

**Example II.28** Given $\mathbf{v} = \begin{bmatrix} -1-4i & 8i & 1 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} 6 & -1 & -1+32i \end{bmatrix}^T$, determine whether $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

**Solution:** $\mathbf{v}^H\mathbf{w} = (-1+4i)(6) - 8i(-1) + 2(0) = -1 + 32i$. Therefore, $\mathbf{v}$ and $\mathbf{w}$ are not orthogonal.

**Example II.29** Given $\mathbf{v} = \begin{bmatrix} -1+2i & 8i & 1 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} 3 & -1 & 3+2i \end{bmatrix}^T$, determine whether $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

**Solution:** $\mathbf{v}^H\mathbf{w} = (-1+2i)(3) + 8i(-1) + 1(3+2i) = 0$. Therefore, $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

## E. PROJECTIONS

Before beginning the discussion of projections, let's first define a *unit vector* and what is meant when we refer to the angle between two vectors. If $\mathbf{u} \in \mathcal{C}^n$ and $\|\mathbf{u}\| = 1$, then $\mathbf{u}$ is a unit vector.

**Definition II.11** Given any vector $\mathbf{x} \in \mathcal{C}^n$, a *unit vector* $\mathbf{u}$ can be formed by dividing the vector $\mathbf{x}$ by its magnitude or 2-norm:

$$\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|}.$$

The proof that shows why $\mathbf{u}$ is a unit vector is straightforward. We start by using the formula for calculating a unit vector, $\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$, then we find the magnitude of $\mathbf{u}$. (If $\mathbf{u}$ is a unit vector, its magnitude must be equal to 1). Happily,

$$\|\mathbf{u}\| = \sqrt{\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right)^T \frac{\mathbf{x}}{\|\mathbf{x}\|}} = \sqrt{\left(\frac{\mathbf{x}^T\mathbf{x}}{\|\mathbf{x}\|\,\|\mathbf{x}\|}\right)} = \sqrt{\left(\frac{\mathbf{x}^T\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\right)} = 1,$$

as anticipated. This argument also holds true for complex vectors. Remember that the magnitude or 2-norm of a vector is equal to the length of the vector. Therefore, when we divide a vector by its magnitude, the result is a vector with unit length, pointing in the direction of the original vector.

**Example II.30** Given $\mathbf{v} = \begin{bmatrix} 4 & 0 & 2 & 4 & -7 \end{bmatrix}^T$, form the unit vector $\mathbf{u}$.

**Solution:**

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{v}}{\sqrt{\mathbf{v}^T\mathbf{v}}} = \frac{1}{\sqrt{89}} \begin{bmatrix} 4 & 0 & 2 & 4 & -7 \end{bmatrix}^T.$$

**Example II.31** Given $\mathbf{v} = \begin{bmatrix} -2+4i & 1+3i & 5 & -3-i \end{bmatrix}^T$, form the unit vector $\mathbf{u}$.

**Solution:**

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{v}}{\sqrt{\mathbf{v}^H\mathbf{v}}} = \frac{1}{\sqrt{65}} \begin{bmatrix} -2+4i & 1+3i & 5 & -3-i \end{bmatrix}^T.$$

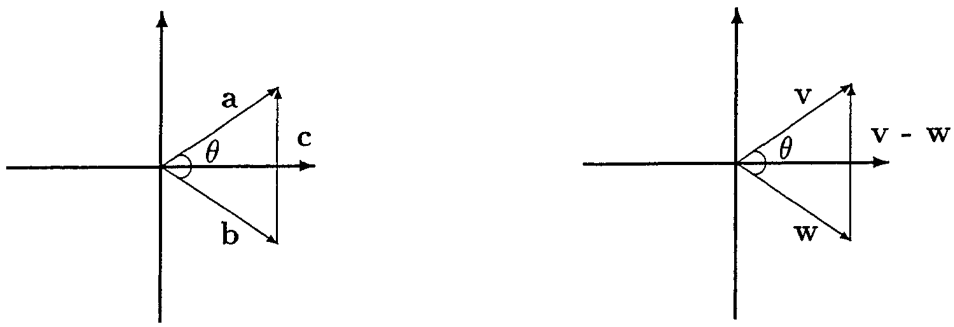


Figure 6. Angle Between Two Vectors

Next, let's define the angle $\theta$ between two vectors in $\mathcal{R}^2$, as shown in Figure 6. Starting in the plane and using the law of cosines, we can find the angle between two arbitrary arrows. Since the arrows in the plane correspond to vectors, all we have to do is use the law of cosines on the vectors and clean up the notation. In Figure 6, the arrows $\mathbf{a}$ and $\mathbf{b}$ are connected by the arrow $\mathbf{a} - \mathbf{b} = \mathbf{c}$. The angle between the arrows $\mathbf{a}$ and $\mathbf{b}$ is $\theta$. The lengths of the arrows are $\|\mathbf{a}\| = a$, $\|\mathbf{b}\| = b$, and $\|\mathbf{c}\| = c$. The law of cosines states $c^2 = a^2 + b^2 - 2ab\cos\theta$. Making the vector substitutions, $\mathbf{a} = \mathbf{v}, \mathbf{b} = \mathbf{w}$, and $\mathbf{c} = \mathbf{v} - \mathbf{w}$ and using the law of cosines, we have

$$\| \mathbf{v} - \mathbf{w} \|^2 = \| \mathbf{v} \|^2 + \| \mathbf{w} \|^2 - 2 \| \mathbf{v} \| \| \mathbf{w} \| \cos\theta.$$

Now expanding $\|\mathbf{v} - \mathbf{w}\|^2$, we have

$$\| \mathbf{v} - \mathbf{w} \|^2 = (\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) = \| \mathbf{v} \|^2 - 2\mathbf{v} \cdot \mathbf{w} + \| \mathbf{w} \|^2.$$

Equating these two equations yields

$$\| \mathbf{v} \|^2 - 2\mathbf{v} \cdot \mathbf{w} + \| \mathbf{w} \|^2 = \| \mathbf{v} \|^2 + \| \mathbf{w} \|^2 - 2 \| \mathbf{v} \| \| \mathbf{w} \| \cos\theta.$$

Canceling like terms leaves us with

$$-2\mathbf{v} \cdot \mathbf{w} = -2 \| \mathbf{v} \| \| \mathbf{w} \| \cos\theta.$$

Finally,

$$\mathbf{v}^H \mathbf{w} = \| \mathbf{v} \| \| \mathbf{w} \| \cos\theta.$$

**Theorem II.3** *Given vectors* $\mathbf{v}$ *and* $\mathbf{w}$ *in* $\mathcal{R}^2$*, the angle* $\theta$ *between* $\mathbf{v}$ *and* $\mathbf{w}$ *is defined by* $\mathbf{v}^H\mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos\theta$*. This generalizes to* $\mathcal{C}^n$*.*

This fact is useful when we are given two vectors and desire to find the angle $\theta$ between them.

**Example II.32** Given $\mathbf{v} = \begin{bmatrix} -2 & 1 & 5 & -3 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} -5 & -4 & 7 & 6 \end{bmatrix}^T$, compute $\theta$.

**Solution:**

$$\cos\theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\| \mathbf{v} \| \| \mathbf{w} \|} = \frac{23}{\sqrt{39}\sqrt{126}}. \text{ Therefore, } \theta = \cos^{-1}\left(\frac{23}{\sqrt{39}\sqrt{126}}\right).$$

Now let's make a couple of observations about the vector $\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ in Figure 7. The length of $\mathbf{v}$ is $\| \mathbf{v} \| = \sqrt{2}$. Also notice the angle that $\mathbf{v}$ makes with the x axis is $\frac{\pi}{4}$ radians or 45°. Applying simple geometry we find that the length of the y coordinate is $\sqrt{2}\sin\left(\frac{\pi}{4}\right) = 1$ and the length of the x coordinate is $\sqrt{2}\cos\left(\frac{\pi}{4}\right) = 1$. This gives us the components of $\mathbf{v}$ along the y and x axes respectively. This can be restated as the projection of $\mathbf{v}$ onto the y and x axes. The projection vectors are represented by $\begin{bmatrix} 0 & 1 \end{bmatrix}^T$ and $\begin{bmatrix} 1 & 0 \end{bmatrix}^T$ respectively. However, we can find the components of $\mathbf{v}$ with respect to any arbitrary vector, not just the coordinate axes.

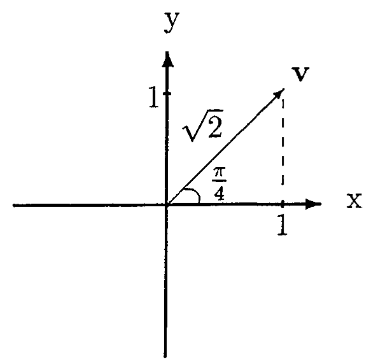


Figure 7. Length of a Vector

**Definition II.12** The *projection of* $\mathbf{v}$ *onto* $\mathbf{w}$, denoted by $proj_{\mathbf{w}}\mathbf{v}$, is the portion of $\mathbf{w}$ which is constructed by passing vector $\mathbf{z}$ perpendicular to $\mathbf{w}$ to the end of $\mathbf{v}$. The $proj_{\mathbf{w}}\mathbf{v}$ extends from the origin to the base of $\mathbf{z}$.

**Definition II.13** Given vectors $\mathbf{v}$, and $\mathbf{w}$, the *magnitude* of $proj_{\mathbf{w}}\mathbf{v}$ is given by

$$\|proj_{\mathbf{w}}\mathbf{v}\| = \|\mathbf{v}\| \cos\theta.$$

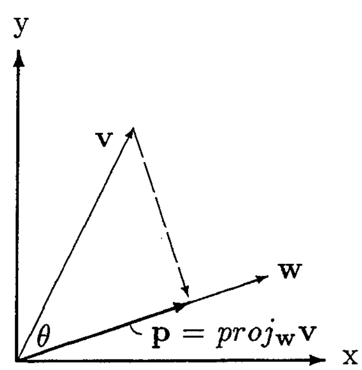


Figure 8. Projection Vector

In Figure 8, we draw a vector $\mathbf{z}$ perpendicular to $\mathbf{w}$ which passes through the end of $\mathbf{v}$. In order to find $proj_{\mathbf{w}}\mathbf{v} = \mathbf{p}$ as in Figure 8, all we have to do is multiply $\|proj_{\mathbf{w}}\mathbf{v}\|$

by a unit vector in the direction of $\mathbf{w}$. To find $\mathbf{p}$, let's first look at our definition in vector form. Replacing $\cos\theta$ in the previous definition with $\frac{\mathbf{w}^T\mathbf{v}}{\|\mathbf{v}\|\|\mathbf{w}\|}$ we have

$$\|proj_{\mathbf{w}}\mathbf{v}\| = \|\mathbf{v}\| \frac{\mathbf{w}^T\mathbf{v}}{\| \mathbf{v} \| \| \mathbf{w} \|}.$$

Canceling like terms we obtain $\|proj_{\mathbf{w}}\mathbf{v}\| = \frac{\mathbf{w}^T\mathbf{v}}{\|\mathbf{w}\|}$. Finally multiply $\| proj_{\mathbf{w}}\mathbf{v} \|$ by a unit vector in the direction of $\mathbf{w}$:

$$\frac{\mathbf{w}}{\|\mathbf{w}\|}\|proj_{\mathbf{w}}\mathbf{v}\| = \frac{\mathbf{w}}{\|\mathbf{w}\|}\frac{\mathbf{w}^T\mathbf{v}}{\|\mathbf{w}\|} = \mathbf{w}\frac{\mathbf{w}^T\mathbf{v}}{\mathbf{w}^T\mathbf{w}} = proj_{\mathbf{w}}\mathbf{v}.$$

**Definition II.14** Given vectors $\mathbf{v}$, and $\mathbf{w}$, the *projection of* $\mathbf{v}$ *onto* $\mathbf{w}$ is

$$proj_{\mathbf{w}}\mathbf{v} = \mathbf{w}\frac{\mathbf{w}^T\mathbf{v}}{\mathbf{w}^T\mathbf{w}} .$$

This gives us the component of $\mathbf{v}$ in the direction of $\mathbf{w}$.

**Example II.33** Given $\mathbf{v} = \begin{bmatrix} -2 & 1 & 5 & -3 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} -5 & -4 & 7 & 6 \end{bmatrix}^T$, compute $proj_{\mathbf{w}}\mathbf{v}$.

**Solution:**

$$proj_{\mathbf{w}}\mathbf{v} = \mathbf{w}\frac{\mathbf{w}^T\mathbf{v}}{\mathbf{w}^T\mathbf{w}} = \begin{bmatrix} -5 \\ -4 \\ 7 \\ 6 \end{bmatrix} \frac{23}{126}.$$

**Example II.34** Given $\mathbf{v} = \begin{bmatrix} 4 & 0 & 2 & 4 & -7 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} -2 & 9 & 6 & -1 & 0 \end{bmatrix}^T$, compute $proj_{\mathbf{w}}\mathbf{v}$.

**Solution:**

$$proj_{\mathbf{w}}\mathbf{v} = \mathbf{w}\frac{\mathbf{w}^T\mathbf{v}}{\mathbf{w}^T\mathbf{w}} = \begin{bmatrix} 4 \\ 0 \\ 2 \\ 4 \\ -7 \end{bmatrix} \frac{0}{122} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Notice this agrees with what we found earlier about the orthogonality of $\mathbf{v}$ and $\mathbf{w}$.

**Example II.35** Given $\mathbf{v} = \begin{bmatrix} -1-4i & 8i & 2 \end{bmatrix}^T$ and $\mathbf{w} = \begin{bmatrix} 6 & -1 & 0 \end{bmatrix}^T$, compute $proj_{\mathbf{w}}\mathbf{v}$.

**Solution:** From a previous example, we know that $\mathbf{v}^H\mathbf{w} = -1 + 32i$. Therefore,

$$proj_{\mathbf{w}}\mathbf{v} = \begin{bmatrix} 6 \\ -1 \\ 0 \end{bmatrix} \frac{-1-32i}{\sqrt{37}}.$$

In the derivation of the definition of projection, we used vector notation. As we have said, there is very little difference between vectors and arrows in the plane. The advantage is that in vector notation we can leave the plane and enter $n$-dimensional spaces and maintain our geometric interpretation. But, as you may have already noticed, we will continue to use the two-dimensional and three-dimensional language even when we are in $n$-dimensional space.

## F. LINEAR COMBINATIONS

Now that we have the ability to perform addition, subtraction, and scalar multiplication on vectors, we can build *weighted sums* of vectors, called *linear combinations.*

**Definition II.15** Given vectors $\mathbf{v}_1, \mathbf{v}_2 \cdots \mathbf{v}_l \in \mathcal{C}^n$, and scalars $\alpha_1, \alpha_2 \cdots \alpha_l$, the weighted sum $\mathbf{w} = \sum_{i=1}^{l} \alpha_i \mathbf{v}_i = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_l\mathbf{v}_l$ is a *linear combination* of the vectors $\mathbf{v}_1, \mathbf{v}_2 \cdots \mathbf{v}_l$ and is also called the vector equation.

**Example II.36** Given $\mathbf{v}_1 = \begin{bmatrix} 1 & 2 \end{bmatrix}^T$ and $\mathbf{v}_2 = \begin{bmatrix} 2 & 4 \end{bmatrix}^T$, find scalars $\alpha_1$ and $\alpha_2$ such that $\mathbf{w} = \begin{bmatrix} 4 & 8 \end{bmatrix}^T$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

**Solution:** Remember, to be a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$, $\mathbf{w}$ must be representable as a weighted sum of vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. Therefore, we must find $\alpha_1$ and $\alpha_2$ such that

$$\begin{bmatrix} 4 \\ 8 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

By inspection, if we let $\alpha_1 = 2$, and $\alpha_2 = 1$, then

$$\begin{bmatrix} 4 \\ 8 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

So we see that there is at least one set of scalars for which $\mathbf{w}$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. Later it will become much clearer how to find these scalars mathematically. The notion of a linear combination is one of the most fundamental ideas in linear algebra and it will come up over and over again throughout this text.

**Example II.37** Given $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, we can form linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, and $\mathbf{v}_4$ as follows:

$$-2\mathbf{v}_1 + 3\mathbf{v}_2 + 7\mathbf{v}_3 - 4\mathbf{v}_4 = \begin{bmatrix} -2 & 3 & 7 & -4 \end{bmatrix}^T \text{ and}$$

$$20\mathbf{v}_1 + -6\mathbf{v}_2 + 0\mathbf{v}_3 + 9\mathbf{v}_4 = \begin{bmatrix} 20 & -6 & 0 & 9 \end{bmatrix}^T.$$

**Example II.38** Given $\mathbf{v}_1 = \begin{bmatrix} i \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ i \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ i \\ 0 \end{bmatrix}$, and $\mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ i \end{bmatrix}$,we can form linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$,and $\mathbf{v}_4$ as follows:

$$-2\mathbf{v}_1 + 3\mathbf{v}_2 + 7\mathbf{v}_3 - 4\mathbf{v}_4 = \begin{bmatrix} -2i & 3i & 7i & -4i \end{bmatrix}^T \text{ and}$$

$$20\mathbf{v}_1 + -6\mathbf{v}_2 + 0\mathbf{v}_3 + 9\mathbf{v}_4 = \begin{bmatrix} 20i & -6i & 0 & 9i \end{bmatrix}^T.$$

## G. LINEAR INDEPENDENCE

We say a set of vectors is *linearly independent* if no vector of the set can be written as a linear combination of the other vectors of the set.

**Definition II.16** Given a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots.., \mathbf{v}_l\}$, a *homogeneous equation* is an equation of the form:

$$\mathbf{0} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_l\mathbf{v}_l.$$

**Definition II.17** If the only solution to the homogeneous equation is the trivial solution, that is $\alpha_1 = \alpha_2 = \cdots = \alpha_l = 0$,then the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots.., \mathbf{v}_l\}$ is said to be *linearly independent.* If a nontrivial solution exists, then the set of vectors is said to be *linearly dependent.*

**Example II.39** Given vectors $\mathbf{w}, \mathbf{v}$, and $\mathbf{u}$, determine whether they are linearly independent.

**Solution:**

$$\mathbf{w} = \begin{bmatrix} 4 \\ 8 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \text{ and } \mathbf{u} = \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix}.$$

By inspection, we see that the following homogeneous equation has a nontrivial solution. Let $\alpha_1 = -1, \alpha_2 = 2$, and $\alpha_3 = 1$.

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = -1 \begin{bmatrix} 4 \\ 8 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Therefore, $\mathbf{w}, \mathbf{v}$, and $\mathbf{u}$ are linearly dependent.

## H. SPANNING SETS

In a previous section we saw how a set of vectors could be used to form new vectors. Here we will use the notion of linear combinations to define the span. The *span* is the set of all possible linear combinations of a given set of vectors. This will also lead to the idea of *vector equations.*

**Definition II.18** Given a set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l\}$, the *span* of $S$, denoted by $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l\}$,or $\text{span}(S)$, is the set of all linear combinations of the vectors in $S$. If $V \subseteq \text{span}(S)$, then $S$ is also known as a *spanning set* of $V$.

**Example II.40** Given $\mathbf{v} = \begin{bmatrix} 0 \\ -1+i \\ 2-3i \\ 3 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 1 \\ 2i \\ 4-7i \\ 0 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} -2 \\ 1 \\ 1 \\ i \end{bmatrix}$, determine whether $\mathbf{b}$ is in $\text{span}\{\mathbf{v}, \mathbf{w}\}$.

**Solution:** In other words, can we find two complex scalars, $\alpha_1$, and $\alpha_2$, such that

$$\alpha_1 \begin{bmatrix} 0 \\ -1+i \\ 2-3i \\ 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 2i \\ 4-7i \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 1 \\ i \end{bmatrix}?$$

In order to find a solution we must solve four equations:

$$\begin{array}{rcrcr} \alpha_1(0) & + & \alpha_2(1) & & -2 \\ \alpha_1(-1+i) & + & \alpha_2(2i) & = & 1 \\ \alpha_1(2-3i) & + & \alpha_2(4-7i) & & 1 \\ \alpha_1(3) & + & \alpha_2(0) & & i. \end{array}$$

However, $\alpha_1$, and $\alpha_2$ must satisfy all four equations simultaneously. By inspection, we see there are no values for $\alpha_1$ and $\alpha_2$ that satisfy all four equations simultaneously. This tells us that $\mathbf{b}$ is not a linear combination of $\mathbf{v}$ and $\mathbf{w}$. Therefore, $\mathbf{b}$ is not in the span of $\mathbf{v}$ and $\mathbf{w}$. Keep this example in mind when we move into systems of linear equations.

## I. BLOCK OPERATIONS

It is often convenient to think of a matrix as having elements that are themselves matrices. By doing this, one can impose structure on a matrix, expose the structure of a matrix, or somehow simplify the way we look at the matrix. For very

large matrices, the concept of block matrices is a must. Matrices whose elements are themselves matrices are called *block matrices.* In a sense, every matrix is a block matrix if you think of its elements as $1 \times 1$ matrices instead of scalars. However, there are many other ways of partitioning a matrix into blocks.

**Example II.41** Given $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$,

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \textit{ and } \mathbf{u}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix},$$

two additional ways of representing A in the form of a block matrix are

$$A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \text{ and } A = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix}.$$

Two block matrices are compatible for addition if they have the same number of rows and columns and the individual blocks are compatibly sized for addition. Two block matrices are compatibly sized for multiplication if the first matrix has as many columns as the second matrix has rows and the number of columns in each column block of the first is the same as the number of rows in the corresponding row block of the second.

**Example II.42** Given $\mathbf{u} = \begin{bmatrix} \mathbf{u}_1^T & \mathbf{u}_2^T & \mathbf{u}_3^T \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}$, compute $\mathbf{u}^T\mathbf{v}$ and $\mathbf{v}\mathbf{u}^T$.

**Solution:**

$$\mathbf{u}^T\mathbf{v} = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{v}_1 & \mathbf{u}_1^T\mathbf{v}_2 & \mathbf{u}_1^T\mathbf{v}_3 \\ \mathbf{u}_2^T\mathbf{v}_1 & \mathbf{u}_2^T\mathbf{v}_2 & \mathbf{u}_2^T\mathbf{v}_3 \\ \mathbf{u}_3^T\mathbf{v}_1 & \mathbf{u}_3^T\mathbf{v}_2 & \mathbf{u}_3^T\mathbf{v}_3 \end{bmatrix}, \text{ and}$$

$$\mathbf{v}\mathbf{u}^T = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1\mathbf{u}_1^T + \mathbf{v}_2\mathbf{u}_2^T + \mathbf{v}_3\mathbf{u}_3^T \end{bmatrix}.$$

The preceding example shows the outer product form of matrix multiplication and the inner-product form of matrix multiplication, respectively. By looking at these two forms, we can see that the inner product form of multiplying two $n \times n$ matrices requires the computation of $n^2$ inner products at a cost of $n$ multiplications and $n-1$ additions each, for a total of $n^3$ multiplications and $n^3 - n^2$ additions. The outer-product form requires the computation of $n$ outer products with $n^2$ multiplications each for $n^3$ multiplications, and $n-1$ additions of $n \times n$ matrices for $n^3 - n^2$ additions. That is, the same number of multiplications and additions are required. However, if we time the two forms of matrix multiplication in MATLAB with two $100 \times 100$ matrices, we will find that it takes roughly 55% longer to do a matrix multiply using the inner product form. There are many other ways of partitioning matrices into blocks that we will see as they become useful.

## J. EXERCISES

1. Given $\mathbf{x} = \begin{bmatrix} -1+2i \\ 2-3i \\ -7i \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} i & -2+4i & 6 \end{bmatrix}$, write or compute

a) $\mathbf{x}^H$ b) $\mathbf{y}^T$ c) $\mathbf{x}+\mathbf{y}$ d) $3\mathbf{x} - 2i\mathbf{y}$

e) $\mathbf{x}+\mathbf{z}$ f) $\mathbf{x}+\mathbf{z}^T$ g) $2\mathbf{z}^H$ h) $\mathbf{z}+\overline{\mathbf{z}}$

2. Given $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$, verify Theorem II.1 parts 2 through 10.

3. Given $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} 5 \\ 2 \\ -2 \end{bmatrix}$, compute

a) $\mathbf{x} \cdot \mathbf{z}$  b) $\mathbf{x} \cdot \mathbf{y}$  c) $\mathbf{x}(\mathbf{z} \cdot \mathbf{y})$  d) $(\mathbf{y} \cdot \mathbf{z})\mathbf{x}$

4. Given $\mathbf{x} = \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} -5 \\ 7 \\ 0 \end{bmatrix}$, compute

a) $\mathbf{x} \cdot \mathbf{z}$  b) $(\mathbf{x} \cdot \mathbf{z})\mathbf{y}$  c) $\mathbf{x}(\mathbf{z} \cdot \mathbf{y})$  d) $(\mathbf{y} \cdot \mathbf{z})\mathbf{x}$

5. Given $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$, verify Theorem II.2.

6. Given $\mathbf{x} = \begin{bmatrix} -5+i \\ 2-3i \\ 1+4i \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -2 \\ 3-i \\ 1+3i \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} -5-2i \\ -4i \\ 3 \end{bmatrix}$, compute

a) $\mathbf{x}^T\mathbf{y}$  b) $\mathbf{y}^T\mathbf{z}$  c) $\mathbf{z}^T\mathbf{x}$  d) $\mathbf{y}^T\mathbf{x}$  e) $\mathbf{z}^T\mathbf{y}$  f) $\mathbf{x}^T\mathbf{z}$

7. Using $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ from question 4, compute

a) $\left(\mathbf{x}^T\mathbf{z}\right)\mathbf{y}$  b) $\mathbf{x}\left(\mathbf{z}^T\mathbf{y}\right)$  c) $\left(\mathbf{y}^T\mathbf{x}\right)\mathbf{z}$  d) $\mathbf{y}\left(\mathbf{x}^T\mathbf{z}\right)$

8. Using through $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ from question 6, compute

a) $\left(\mathbf{x}^H\mathbf{z}\right)\mathbf{y}$  b) $\mathbf{x}\left(\mathbf{z}^H\mathbf{y}\right)$  c) $\left(\mathbf{y}^H\mathbf{x}\right)\mathbf{z}$  d) $\mathbf{y}\left(\mathbf{x}^H\mathbf{z}\right)$

9. Given $\mathbf{w} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 4 \\ -2 \\ 2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} 2 \\ -2 \\ -6 \end{bmatrix}$, find all the orthogonal pairs.

10. Given $\mathbf{x} = \begin{bmatrix} 4 \\ -2 \\ 0 \\ 2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -3 \\ 4 \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} 1 \\ 2 \\ 5 \\ 0 \end{bmatrix}$,

a) Compute $\|\mathbf{x}\|$.

b) Compute $\|\mathbf{y}\|$.

c) Compute $\|\mathbf{z}\|$.

d) Are $\mathbf{x}$ and $\mathbf{y}$ orthogonal?

e) Are $\mathbf{z}$ and $\mathbf{y}$ orthogonal?

f) Are $\mathbf{x}$ and $\mathbf{z}$ orthogonal?

11. Given $\mathbf{x} = \begin{bmatrix} 4+2i \\ -2 \\ -3+i \\ -5i \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -2i \\ 1+i \\ -3i \\ -1+10i \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} 1 \\ 2i \\ -1+4i \\ i \end{bmatrix}$,

a) Compute $\|\mathbf{x}\|$.

b) Compute $\|\mathbf{y}\|$.

c) Compute $\|\mathbf{z}\|$.

d) Are $\mathbf{x}$ and $\mathbf{y}$ orthogonal?

e) Are $\mathbf{z}$ and $\mathbf{y}$ orthogonal?

f) Are $\mathbf{x}$ and $\mathbf{z}$ orthogonal?

12. Given $\mathbf{v} = \begin{bmatrix} -4 \\ 2 \\ -5 \\ 7 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} -16 \\ 6 \\ 3 \\ -2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -i \\ 2-i \\ -3+2i \\ 6+5i \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 6 \\ -3-4i \\ -5i \\ 2 \end{bmatrix}$,

compute

a) $\|\mathbf{v}\|$ b) $\|\mathbf{w}\|$ c) $\|\mathbf{x}\|$ d) $\|\mathbf{y}\|$ e) Find all the orthogonal pairs.

13. Given $\mathbf{v} = \begin{bmatrix} -2 \\ 0 \\ 3 \\ 1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 3 \\ -5 \\ -4 \\ 2 \end{bmatrix}$, compute

a) $\mathbf{w}^T\mathbf{v}$

b) $\mathbf{w}^T\mathbf{w}$

c) $\cos\theta$, where $\theta$ is the angle between $\mathbf{v}$ and $\mathbf{w}$.

d) $\text{proj}_{\mathbf{w}}\mathbf{v}$

14. Given $\mathbf{v} = \begin{bmatrix} -2 \\ 3 \\ 3 \\ -2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 3 \\ -5 \\ -4 \\ 0 \end{bmatrix}$, compute

a) $\mathbf{w}^T\mathbf{v}$

b) $\mathbf{w}^T\mathbf{w}$

c) $\cos\theta$, where $\theta$ is the angle between $\mathbf{v}$ and $\mathbf{w}$.

d) $\text{proj}_{\mathbf{w}}\mathbf{v}$

15. Given $\mathbf{v} = \begin{bmatrix} -2i \\ 3 \\ 3-i \\ -2+4i \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 3+i \\ -5i \\ -4 \\ 0 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -2 \\ 0 \\ 3 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 3 \\ -5 \\ -4 \\ 2 \end{bmatrix}$, compute $\cos\theta$, where $\theta$ is the angle between

a) $\mathbf{w}$ and $\mathbf{v}$ b) $\mathbf{w}$ and $\mathbf{x}$ c) $\mathbf{x}$ and $\mathbf{y}$ d) $\mathbf{v}$ and $\mathbf{y}$ e) $\mathbf{x}$ and $\mathbf{v}$

f) $\mathbf{w}$ and $\mathbf{y}$

16. Given vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x} \in \mathcal{R}^3$, and the vector equation $-2(2\mathbf{x} + \mathbf{v}) = 3(\mathbf{w} - \mathbf{u})$, express $\mathbf{x}$ as a linear combination of $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$.

17. Given $\mathbf{v} = \begin{bmatrix} -2i \\ 3 \\ 3-i \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} -2i \\ i \\ 2+2i \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 2 \\ -i \\ -3i \end{bmatrix}$

a) Write a vector equation with $\mathbf{w}, \mathbf{x}$, and $\mathbf{y}$ as independent variables multiplied by some scalar, and $\mathbf{v}$ as the dependent variable.

b) Find $\mathbf{w} + \mathbf{x} + \mathbf{y}$. Is $\mathbf{v}$ in the span $\{\mathbf{w}, \mathbf{x}, \mathbf{y}\}$?

c) Are $\mathbf{v}, \mathbf{w}, \mathbf{x}$, and $\mathbf{y}$ linearly independent?

18. Given $A = \begin{bmatrix} 1 & -3 & -2 \\ -2 & 2 & 6 \\ 2 & -8 & -3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 6 \\ -3 \\ 5 \end{bmatrix}$, is $\mathbf{b}$ a linear combination of the columns of A?

19. Given $A = \begin{bmatrix} 2 & -1 & 3 & 4 \\ 1 & 0 & -2 & 7 \\ 3 & -3 & 1 & 5 \\ 2 & 1 & 4 & 4 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 9 \\ 11 \\ 8 \\ 10 \end{bmatrix}$, is $\mathbf{b}$ a linear combination of the columns of A?

# III. LINEAR TRANSFORMATIONS

*Linear transformations* are extremely powerful tools. In reality, a linear transformation is simply a way of looking at what goes into, and what comes out of a black box. When speaking about functions of the form $y = f(x)$, we take a variable $x$, perform some function on it, and get a result $y$. The function is the black box, with input $x$ and output $y$. We use special terms to define ideas about this system. The function $f(x)$ is a "black box", which assigns $y$ (an element from the range) to $x$ (an element from the domain). In other words, $x$ is transformed by the function $f(x)$ into $y$. When discussing linear transformations involving vectors and matrices, we will use special terms to define similar ideas. Recall that a *function* is a rule that assigns to each element in a set **A** (the domain), one and only one element from a set **B** (the codomain). When dealing with vectors and matrices we will use the terms *map* or *transform* instead of assign. Therefore a linear transformation ("black box") will map (or transform) a vector $\mathbf{x}$ to a vector $\mathbf{b}$. Given a function $f$, the notation $f : \mathcal{A} \longrightarrow \mathcal{B}$ means that $f$ is a transformation from $\mathcal{A}$ to $\mathcal{B}$.

Now let's make the jump to linear algebra and see that transformations are functions.

**Definition III.1** A *transformation* from $\mathcal{R}^n$ to $\mathcal{R}^m$ is a rule that assigns to each vector $\mathbf{x} \in \mathcal{R}^n$ a vector $\mathbf{y} \in \mathcal{R}^m$. Formally, we say that a transformation $T$, denoted by $T : \mathcal{R}^n \longrightarrow \mathcal{R}^m$, maps $\mathcal{R}^n$ into $\mathcal{R}^m$.

We can use just about any rule to define a transformation. In matrix algebra, though, we are particularly interested in transformations of the form $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$. Such a transformation takes a vector $\mathbf{x}$, and transforms it by multiplying on the left by a compatibly-sized matrix A. Since we know that a vector $\mathbf{x}$ multiplied by a matrix A is some vector $\mathbf{b}$, then we can write $\mathrm{A}\mathbf{x} = \mathbf{b}$, or in other words, a matrix A transforms a vector $\mathbf{x}$ into a vector $\mathbf{b}$. Suppose that $\mathbf{x} \in \mathcal{R}^n$ is mapped to a vector $\mathbf{b} \in \mathcal{R}^m$ by the matrix A. Based on our knowledge of matrix multiplication, what can we say about

the dimensions of A? Since $\mathbf{x} \in \mathcal{R}^n$ and $\mathbf{b} \in \mathcal{R}^m$ then $\mathrm{A} \in \mathcal{R}^{m \times n}$. Now that we know what a transformation is, let's define exactly what is meant by *linear* transformation.

**Definition III.2** $T : \mathcal{R}^n \longrightarrow \mathcal{R}^m$ is a *linear transformation* if, for any vectors $\mathbf{x}$ and $\mathbf{y} \in \mathcal{R}^n$ and any scalar $c \in \mathcal{R}$, the following two conditions hold:

1) $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$
2) $T(c\mathbf{x}) = cT(\mathbf{x})$.

This definition implies that it doesn't matter whether addition or scalar multiplication is performed before or after the transformation; the same results will be achieved. In order to clarify this idea, we will use the analogy of a chef in the kitchen preparing a salad and baking bread. Generally, the act of chopping is a linear transformation. You can either chop the vegetables up individually, then mix them, or you can mix the whole vegetables and then chop the entire mixture. Either way, you will end up with a bowl of chopped vegetables. However, baking in general is not a linear transformation. A good chef will follow the baking instructions and mix the ingredients first, then bake the bread. A skeptical mathematician will bake the ingredients individually and then try to mix the ingredients. Surely, bread can't be made this way. The order of baking and mixing is very important.

**Example III.1** Given $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$, determine whether $T$ is a linear transformation.

**Solution:** In order to determine this, $T$ must satisfy the two conditions previously stated. We start with the first condition, $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$.
Working with the left side of the equality by definition of $T$, we have

$$T(\mathbf{x}+\mathbf{y}) = \mathrm{A}(\mathbf{x}+\mathbf{y}).$$

Since matrix multiplication is distributive, it follows that

$$\mathrm{A}(\mathbf{x}+\mathbf{y}) = \mathrm{A}\mathbf{x} + \mathrm{A}\mathbf{y}.$$

Now, working with the right-hand side of the equality,

$$T(\mathbf{x}) + T(\mathbf{y}) = \mathrm{A}\mathbf{x} + \mathrm{A}\mathbf{y}.$$

Therefore,

$$T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}).$$

This satisfies the first condition, but what about the second condition? We must show that

$$T(c\mathbf{x}) = cT(\mathbf{x}).$$

Working with the left-hand side of the equality, we have

$$T(c\mathbf{x}) = \mathrm{A}(c\mathbf{x}).$$

Since scalar multiplication of vectors and matrices is commutative, then

$$\mathrm{A}(c\mathbf{x}) = c\mathrm{A}\mathbf{x}.$$

Now working with the right-hand side of the equality, we have

$$cT(\mathbf{x}) = c\mathrm{A}\mathbf{x}.$$

Therefore, $T(c\mathbf{x}) = cT(\mathbf{x})$, which satisfies the second condition. Since conditions one and two are satisfied, $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ is a linear transformation. If either condition did not hold, then the transformation would not have been linear.

**Example III.2** Given $T(\mathbf{x}) = \begin{bmatrix} \sin(x_1) \\ \sqrt{x_2} \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, determine whether $T$ is a linear transformation.

**Solution:** Again starting with condition one,

$$T(\mathbf{x} + \mathbf{y}) = T\left(\begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}\right) = \begin{bmatrix} \sin(x_1 + y_1) \\ \sqrt{x_2 + y_2} \end{bmatrix}, \text{ while}$$

$$T(\mathbf{x})+T(\mathbf{y}) = \begin{bmatrix} \sin(x_1) \\ \sqrt{x_2} \end{bmatrix} + \begin{bmatrix} \sin(y_1) \\ \sqrt{y_2} \end{bmatrix} = \begin{bmatrix} \sin(x_1) & + & \sin(y_1) \\ \sqrt{x_2} & + & \sqrt{y_2} \end{bmatrix}, \text{ so}$$

$$T(\mathbf{x}+\mathbf{y}) = \begin{bmatrix} \sin(x_1+y_1) \\ \sqrt{x_2+y_2} \end{bmatrix} \neq \begin{bmatrix} \sin(x_1) & + & \sin(y_1) \\ \sqrt{x_2} & + & \sqrt{y_2} \end{bmatrix} = T(\mathbf{x})+T(\mathbf{y}).$$

Because condition one does not hold, $T$ is not a linear transformation.

**Example III.3** Given $T(\mathbf{x}) = \begin{bmatrix} 2x_1 \\ x_2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, determine whether $T$ is a linear transformation.

**Solution:** Again starting with condition one, we have

$$T(\mathbf{x}+\mathbf{y}) = T\left(\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \end{bmatrix}\right) = \begin{bmatrix} 2(x_1+y_1) \\ x_2+y_2 \end{bmatrix}, \text{ and}$$

$$T(\mathbf{x})+T(\mathbf{y}) = \begin{bmatrix} 2x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2(x_1+y_1) \\ x_2+y_2 \end{bmatrix}.$$

Condition one holds, but we must still check condition two:

$$T(c\mathbf{x}) = T\left(c\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} c2x_1 \\ cx_1 \end{bmatrix} = c\begin{bmatrix} 2x_1 \\ x_1 \end{bmatrix}, \text{ and}$$

$$cT(\mathbf{x}) = cT\left(\begin{bmatrix} x_1 \\ x_1 \end{bmatrix}\right) = c\begin{bmatrix} 2x_1 \\ x_1 \end{bmatrix}.$$

Because condition two also holds, we say that $T$ is a linear transformation.

In engineering and physics, one of the most useful tools is the simplification of a problem. By splitting a problem up into smaller problems, large, seemingly unsolvable problems become a number of smaller solvable problems. The solutions to the smaller problems can then be combined to get a solution to the larger problem.

This process is based on the *superposition principle* and is a generalization of our two conditions:

$$T(c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_n\mathrm{x}_n) = c_1T(\mathrm{x}_1) + c_2T(\mathrm{x}_2) + \cdots + c_nT(\mathrm{x}_n).$$

## A. SPECIAL MATRICES

There are a number of matrices that, because of their frequency of use and importance, have been given special names. These *special matrices* have specific useful properties. We start with perhaps the most important special matrix of all.

### 1. The Identity Matrix

The first special matrix, and probably the most widely used, is the *identity* matrix. The identity matrix performs the same function as the number 1 in multiplication; $1a = a$ and $b1 = b$. This matrix is denoted by I, and has ones along the *main diagonal* and zeros elsewhere. (The main diagonal elements of a matrix A are those elements that are indexed by $\mathrm{a}_{ii}$.)

**Example III.4** The matrix $\mathrm{I} \in \mathcal{R}^{4\times 4}$ is an identity matrix.

$$\mathrm{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

When dealing with square matrices, the identity matrix is special because IA = AI = A. This is a special case in which matrix multiplication commutes.

**Example III.5** Given $\mathrm{A} = \begin{bmatrix} 3 & i \\ -1 & 8 \end{bmatrix}$, and the identity matrix $\mathrm{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, compute IA and AI.

**Solution:**

$$\text{IA} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & i \\ -1 & 8 \end{bmatrix} = \begin{bmatrix} 1\cdot 3 + 0(-1) & 1\cdot i + 0\cdot 8 \\ 0\cdot 3 + 1(-1) & 0\cdot i + 1\cdot 8 \end{bmatrix} = \begin{bmatrix} 3 & i \\ -1 & 8 \end{bmatrix} \text{ and}$$

$$\text{AI} = \begin{bmatrix} 3 & i \\ -1 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3\cdot 1 + i\cdot 0 & 3\cdot 0 + i\cdot 1 \\ -1(1) + 8\cdot 0 & -1(0) + 8\cdot 1 \end{bmatrix} = \begin{bmatrix} 3 & i \\ -1 & 8 \end{bmatrix}.$$

Notice that IA = AI. In most cases, the dimensions of the identity matrix are inferred from the context of the associated matrices.

## 2. Permutation Matrices

A second very important special matrix is a *permutation* matrix. This square matrix is denoted by P, and is very closely related to the identity matrix. PA transforms A by exchanging the rows of A. AP exchanges the columns of A. The permutation matrix transforms a matrix from $\mathcal{R}^{m\times n}$ to $\mathcal{R}^{m\times n}$ and a vector from $\mathcal{R}^n$ to $\mathcal{R}^n$.

**Example III.6** Given $\text{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\text{A} = \begin{bmatrix} 2 & 6i \\ 1 & -3 \end{bmatrix}$, compute PA.

**Solution:**

$$\text{PA} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 6i \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 2 & 6i \end{bmatrix}.$$

P exchanges the rows of A.

**Example III.7** Given $\text{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ and $\text{A} = \begin{bmatrix} 5 & -10+8i & 0 \\ -5i & 2 & 6i \\ 1-i & 1 & -3 \end{bmatrix}$, compute AP.

**Solution:**

$$\text{AP} = \begin{bmatrix} 5 & -10+8i & 0 \\ -5i & 2 & 6i \\ 1-i & 1 & -3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 & -10+8i \\ -5i & 6i & 2 \\ 1-i & -3 & 1 \end{bmatrix}.$$

P exchanges the columns of A.

An important concept to grasp is that certain rows of matrix A can remain in the same position when computing PA. This is done by placing a one in any such row's main diagonal position of matrix P. Additionally, certain columns of A can remain in the same position when computing AP. This is done by placing a one in any such column's main diagonal position of matrix P.

**Example III.8** Given P $= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ and A $= \begin{bmatrix} 0 & 3 & 4 \\ -5 & 6 & 1 \\ i & 2i & 3 \end{bmatrix}$, compute matrix PA.

**Solution:**

$$\text{PA} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 3 & 4 \\ -5 & 6 & 1 \\ i & 2i & 3 \end{bmatrix} = \begin{bmatrix} i & 2i & 3 \\ -5 & 6 & 1 \\ 0 & 3 & 4 \end{bmatrix}.$$

Since the 1 in the second row of P occupies the main diagonal position, row two of A remains fixed, while rows one and three are interchanged.

## 3. Orthogonal Projection Matrices

An *orthogonal projection* matrix can find the components of a vector **v** along any given axis. The orthogonal projection matrix transforms a vector from $\mathcal{R}^n$ to $\mathcal{R}^n$.

**Example III.9** Given Q $= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and **v** $= \begin{bmatrix} x & y \end{bmatrix}^T$, compute Q**v**.

**Solution:** Q**v** is the component of **v** which lies on the x axis:

$$\text{Q}\mathbf{v} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}.$$

Q takes **v** and projects it onto the $x$ axis.

**Example III.10** Given $Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} x & y & z \end{bmatrix}^T$, compute Q**v**.

**Solution:**

$$Q\mathbf{v} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}.$$

Q projects **v** onto the $(x, y)$ plane. If we continue with a $4 \times 4$ orthogonal projection matrix, we could see what a four-dimensional object would look like if we projected the four-dimensional object into three dimensions. This is the same idea of drawing a three-dimensional box on a two-dimensional piece of paper.

### 4. Rotation Matrices

A *rotation* matrix transforms a vector **v** by rotating it about the origin through an angle of $\theta$ degrees. Its derivation goes beyond the scope of this text, but its utility merits an example. A rotation matrix transforms a vector from $\mathcal{R}^n$ to $\mathcal{R}^n$.

**Example III.11** Given $T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$, and $\theta = \frac{\pi}{4}$, compute T**v**.

**Solution:**

$$T\mathbf{v} = \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}.$$

T applied to **v** rotates **v** about the origin counter-clockwise by $\frac{\pi}{4}$ radians, or 45°.

## B. MATRIX NORMS

Although there are many different matrix norms, we are interested in a family of norms called $p$-norms. When we found $||\mathbf{b}||$, we called it the length. The next question is, what is $||A||$? This is not an easy question to answer. First let us look at A**x**. When one forms the product A**x**, where $A \in \mathcal{C}^{n\times m}$ and $\mathbf{x} \in \mathcal{C}^{m\times 1}$, the result is a

vector $\mathbf{b} \in \mathcal{C}^{n\times 1}$ or $A\mathbf{x} = \mathbf{b}$. In some sense, A transforms the $\mathbf{x}$ into a new $\mathbf{b}$ that is a linear combination of the columns of A. It is useful to know how the norm of $A\mathbf{x}$, denoted $\|A\mathbf{x}\|$, relates to the norm of $\mathbf{x}$, denoted $\|\mathbf{x}\|$.

**Definition III.3** Given $\mathbf{x} \in \mathcal{R}^m$, and $A \in \mathcal{R}^{n\times m}$, the *p-norm* of A, denoted by $\|A\|_p$, is defined by $\max\limits_{x\neq 0} \|A\|_p = \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p}$, where $p$ represents a specific vector norm.

The most common of these norms are the 1, 2, and $\infty$ norms, denoted by $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$, respectively. In effect, the norm of a matrix tells us just how much a matrix can stretch vectors. If $\|A\| < 1$ then we know that the matrix A can transforms vectors $\mathbf{x}$ into shorter vectors $\mathbf{b}$, i.e., A shrinks vectors. If $\|A\| > 1$, then we know the matrix A transforms some vectors $\mathbf{x}$ into longer vectors $\mathbf{b}$, i.e., A stretches some vectors. The calculation of the matrix norm varies depending on which norm is desired. The easiest $p$-norm to find is the $\infty$- norm:

$$\|A\|_\infty = \max_{1\leq i\leq n} \left\{ \sum_{k=1}^{n} | a_{ik} | \right\}$$

or equivalently, the maximum absolute row sum. The 1-norm is also easy to calculate.

$$\|A\|_1 = \max_{1\leq j\leq n} \left\{ \sum_{k=1}^{n} | a_{kj} | \right\}$$

or equivalently, the maximum absolute column sum.

**Example III.12** Given $A = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$, compute $\|A\|_1$ and $\|A\|_\infty$.

**Solution:** $\|A\|_1$ is the maximum absolute column sum.

$$\|A\|_1 = \max\{|-2| + |1| + 0 = 3,\ \ |1| + |-2| + |1| = 4,\ \ 0 + |1| + |-2| = 3\}.$$

The maximum value of the absolute column sum is 4, therefore, $\|A\|_1 = 4$. This result indicates that $A\mathbf{x}$ might stretch $\mathbf{x}$ to as much as 4 times its length. The $\infty$-norm is the maximum absolute row sum:

$$\|A\|_\infty = \max\{|-2| + |1| + 0 = 3, |1| + |-2| + |1| = 4, 0 + |1| + |-2| = 3\}.$$

The maximum absolute row sum is 4, therefore, $\|A\|_{\infty} = 4$. This result also indicates that $A\mathbf{x}$ might stretch $\mathbf{x}$.

Due to the difficulty in computing the 2-norm, we leave it for another class.

## C. THE RANGE AND NULL SPACE OF A MATRIX

In this section we introduce two important concepts in linear algebra.

**Definition III.4** Given a matrix $A \in \mathcal{C}^{n\times m}$, the *range of* A, denoted by R(A), is given by $R(A) = \{\mathbf{y} \in \mathcal{C}^n \mid \mathbf{y} = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathcal{C}^m\}$.

In effect, R(A) is the set of all linear combinations of the columns of A, and so is sometimes called the *column space*. The *null space* is related to the column space. Using the columns of A, we form a homogeneous equation,

$$x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \end{bmatrix} + x_3 \begin{bmatrix} a_{1,3} \\ a_{2,3} \\ a_{3,3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Any $\mathbf{x}$ that satisfies the homogeneous equation is in N(A).

**Definition III.5** Given a matrix $A \in \mathcal{C}^{n\times m}$, the *null space* of A, denoted by N(A), is given by $N(A) = \{\mathbf{x} \in \mathcal{C}^n \mid A\mathbf{x} = \mathbf{0}\}$.

**Example III.13** Given $A = \begin{bmatrix} -2 & 6 & -4 \\ 5 & -3 & -2 \end{bmatrix}$, compute N(A).

**Solution:** Form the homogeneous vector equation:

$$x_1 \begin{bmatrix} -2 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 6 \\ -3 \end{bmatrix} + x_3 \begin{bmatrix} -4 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

By inspection, we see that any $\mathbf{x} = \begin{bmatrix} t & t & t \end{bmatrix}^T$, where $t \in \mathcal{R}$, is transformed into the zero vector.

## D. RANK

We have already defined linear independence of vectors. In this section we will think of the columns of a matrix as individual vectors, and then determine the linear independence or linear dependence of the columns. This enables us to determine the *rank* of a matrix. We start this section with an example, to review what is meant when we say that the columns of a matrix are linearly independent or dependent.

**Example III.14** Given $A = \begin{bmatrix} -1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$, determine whether the columns of A are linearly independent.

**Solution:** We start by letting the columns of the matrix be the individual vectors

$$\mathbf{a}_1 = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \text{ and } \mathbf{a}_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Now we ask whether vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, and $\mathbf{a}_4$ are linearly independent or dependent. Using what we learned earlier, we form a linear combination of the vectors into a homogeneous equation $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 + c_4\mathbf{a}_4 = \mathbf{0}$. If the only solution to this equation is the trivial solution, scalars $c_1, c_2, c_3$, and $c_4 = 0$, then the vectors are linearly independent. If a nontrivial solution exists, then the vectors are linearly dependent. At this time we will determine the values for $\mathbf{c}$ by inspection, and ignore the methods for computing them. We leave that to later sections. If we rearrange the homogeneous equation into $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 = -c_4\mathbf{a}_4$ and set all of the constants to 1, we have $\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 = -\mathbf{a}_4$. Substituting the vectors into the homogeneous equation yields

$$\begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \end{bmatrix}.$$

Since $-\mathbf{a}_4$ can be written as a linear combination of $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$, the columns of A are linearly dependent. This looks like the homogeneous matrix equation, except that we have moved the fourth column to the right-hand side of the equation. We could have instead moved the first column to the right-hand side of the equality.

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Again we see that the columns of A are linearly dependent, because a nontrivial solution to the homogeneous equation exists: $c_1 = c_2 = c_3 = c_4 = 1$. Since the columns of A can be thought of as vectors, and since the vectors are linearly dependent, then the columns of A are linear dependent. Notice that three vectors are necessary to form a linear combination of the fourth vector. This is not a coincidence. If we take any three columns of A and form the homogeneous equation $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 = \mathbf{0}$, we will find that the three columns are linearly independent.

Now that we have reviewed what is meant by linear dependence and independence, we can move on to the idea of rank.

**Definition III.6** The *column rank* of a matrix is equal to the number of linearly independent columns. The *row rank* of a matrix is equal to the number of linearly independent rows. The column rank is equal to the row rank and is called the *rank* of a matrix. If the number of linearly independent columns or rows is equal to the smallest dimension of the matrix, then the matrix is said to have *full rank.*

**Theorem III.1** *Given* $\mathrm{A} \in \mathcal{R}^{n \times m}, \mathrm{x} \in \mathcal{R}$, *and* $T(x) = \mathrm{Ax}$, *then the following are true:*

a) $\text{rank}(A) = r$

b) $r \leq n$ and $r \leq m$

c) A has $r$ linearly independent columns

d) A has $r$ linearly independent rows

e) $\text{rank}(A) = \text{rank}(A^T)$

f) The linearly independent columns of A span $R(A)$

g) The linearly independent rows of A span $R(A^T)$

**Example III.15** Given $A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$, determine the number of linearly independent rows of A.

**Solution:** Notice that the fourth column is a linear combination of the first three columns; i.e. col4 = col1 + col2 + col3. Therefore, A has three linearly independent columns. This implies that $\text{rank}(A) = 3$, and so A also has three linearly independent rows.

**Example III.16** Given $A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$, determine the number of linearly independent columns of A.

**Solution:** Notice that the third row is a linear combination of the first two rows; i.e. row3 = row1 + row2. Also, the fourth row is a linear combination of the first two rows; i.e. row4 = row1 − row2. Therefore, A has two linearly independent rows. This implies that $\text{rank}(A) = 2$, and so A also has two linearly independent columns.

**Example III.17** Given $A = \begin{bmatrix} 1 & 5 & 9 & 13 \\ 2 & 6 & 10 & 14 \\ 3 & 7 & 11 & 15 \\ 4 & 8 & 12 & 16 \end{bmatrix}$, determine $\text{rank}(A)$.

**Solution:** In this example, it is doubtful whether the reader could come up with a solution by inspection. Other methods are required to solve this problem. Therefore, at some point, we must shift our attention from the concept of linear independence to the algebraic details of finding a nontrivial solution to the homogeneous equation. We promise to discuss these details later.

## E. ELEMENTARY MATRICES

*Elementary matrices are* a family of special matrices that can be used to perform specific operations on matrices. Elementary matrices will enable us to find new ways to solve systems of linear equations. Each elementary matrix E performs one row operation on a matrix. This operation is known as an *elementary row operation.* There are three elementary row operations: *row interchange*, *row replacement,* and *scalar multiplication of a row.* We form an elementary matrix by applying one of the row operations to the identity matrix.

### 1. Row Interchange

A row interchange, as its name suggests, occurs when any two rows of a matrix are interchanged.

**Example III.18** Given $\mathrm{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, form an elementary matrix which interchanges the second and fourth rows of a matrix.

**Solution:** Interchange the second and fourth rows of I to form the elementary matrix

$$\mathrm{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Using our newly formed elementary matrix E, we can now interchange the second and fourth rows of any $4 \times m$ matrix by left-multiplying that matrix by E. As we learned previously, E is a permutation matrix.

**Example III.19** Given $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$ and $E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$, compute EA.

**Solution:**

$$EA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 13 & 14 & 15 & 16 \\ 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 \end{bmatrix}.$$

Notice that E interchanges the second and fourth rows of A.

**Example III.20** Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, compute EA.

**Solution:**

$$EA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}.$$

Notice that E interchanges the two rows of A.

## 2. Scalar Multiplication of a Row

This operation replaces a row of a matrix with a scalar multiple of the row.

**Example III.21** Given $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$ and $E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, compute EA.

**Solution:**

$$EA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ -10 & -12 & -14 & -16 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}.$$

Notice that E multiplies the second row of A by $-2$.

**Example III.22** Given $A = \begin{bmatrix} 3 & 5 \\ -2 & 7 \end{bmatrix}$ and $E = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$, compute EA.

**Solution:**

$$EA = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ -2 & 7 \end{bmatrix} = \begin{bmatrix} 12 & 20 \\ -2 & 7 \end{bmatrix}.$$

Notice that E multiples the first row of A by 4.

## 3. Row Replacement

This operation will add a scalar multiple of one row to another row. Again we begin with an identity matrix.

**Example III.23** Given $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, form an elementary matrix which adds twice the third row to the first row of I.

**Solution:**

$$E = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

**Example III.24** Given $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ and $E = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, compute EA.

**Solution:**

$$\text{EA} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 15 & 18 & 21 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Notice that E adds twice the third row to the first row of A.

In each of the previous examples, the elementary matrices performed one and only one elementary operation. However, we can multiply many elementary matrices to form a single matrix, A, that will perform those same elementary row operations at one time.

**Example III.25** Given $\text{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$ and $\text{E} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & -2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$, compute EA.

**Solution:**

$$\text{EA} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & -2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} = \begin{bmatrix} 13 & 14 & 15 & 16 \\ -10 & -12 & -14 & -16 \\ 8 & 8 & 8 & 8 \\ 1 & 2 & 3 & 4 \end{bmatrix}.$$

In this example, E actually performs three elementary row operations.

## F. EXERCISES

1. Given $T : \mathcal{R}^2 \longrightarrow \mathcal{R}^3$ and $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ 0 \\ (x_1 + x_2)^2 \end{bmatrix}$, is $T$ a linear transformation?

2. Given $T : \mathcal{R}^2 \longrightarrow \mathcal{R}^3$ and $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 \\ 2x_1 - x_2 \\ x_2 \end{bmatrix}$, is $T$ a linear transformation?

3. Given $T : \mathcal{R}^2 \longrightarrow \mathcal{R}^3$ and $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + 4x_2 \\ x_1 x_2 \end{bmatrix}$, is $T$ a linear transformation?

4. Given the following matrices, compute the 1-norms.

a) $\mathrm{A} = \begin{bmatrix} 1 & -3 & 2 \\ 7 & 3 & -2 \\ 5 & 8 & 0 \end{bmatrix}$ b) $\mathrm{B} = \begin{bmatrix} 0 & 3 & 4 \\ 8 & 9 & 1 \\ 6 & 0 & 0 \end{bmatrix}$ c) $\mathrm{C} = \begin{bmatrix} 1 & -2 & 4 \\ 2 & 4 & 0 \\ -1 & 7 & -2 \end{bmatrix}$

5. Given the following matrices, compute the $\infty$-norms.

a) $\mathrm{A} = \begin{bmatrix} 1 & -3 & 2 \\ 7 & 3 & -2 \\ 5 & 8 & 0 \end{bmatrix}$ b) $\mathrm{B} = \begin{bmatrix} 0 & 3 & 4 \\ 8 & 9 & 1 \\ 6 & 0 & 0 \end{bmatrix}$ c) $\mathrm{C} = \begin{bmatrix} 1 & -2 & 4 \\ 2 & 4 & 0 \\ -1 & 7 & -2 \end{bmatrix}$

6. Given the following matrices, determine whether they are elementary matrices or not. If they are elementary matrices, state which operation they perform on a compatibly-sized matrix.

a) $\mathrm{E}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}$ b) $\mathrm{E}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ c) $\mathrm{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

d) $E_4 = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ e) $E_5 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ f) $E_6 = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

7. Given $E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $E_2 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

a) What elementary row operations do the elementary matrices $E_1$,$E_2$, and $E_3$ perform?

b) Compute $E_1E_2E_3$.

c) Compute $E_3E_2E_1$.

d) Are the solutions to parts b and c equivalent? Why or why not?

8. Given $A = \begin{bmatrix} -3 & 2 & 1 & 0 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$,

a) Are the columns linearly independent, or linearly dependent?

b) Are the rows linearly independent, or linearly dependent?

c) How many linearly independent columns in A?

d) How many linearly independent rows in A?

e) What is the rank(A)?

9. Given $A = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$,

a) What is the rank(A)?

b) Compute the solution set $\mathbf{x}$ to the equation $A\mathbf{x} = 0$.

# IV. SYSTEMS OF LINEAR EQUATIONS AND SOLVING LINEAR SYSTEMS

The first step in solving any real-world problem mathematically is defining the variables and the relationships between the variables. The variables might be, for example, time, distance, speed, costs, materials or profit. The relationships might be known formulas such as $velocity = \frac{distance}{time}$. Once we have defined the variables and we have formed equations using the relationships which describe the problem, we can turn to mathematics to solve the problem. The problems we will be dealing with in this text are linear systems. There are many ways to solve systems of linear equations. One method for solving them is the *method of substitution.* This is the most fundamental of all the methods. It requires that we solve for a single variable in one equation and then substitute the result into the other equations. This is a standard method for two or three variables, but for real-world problems with more then three variables substitution is, at best, cumbersome. In this we will look at three methods for solving linear systems: *Gaussian elimination, matrix inversion,* and LU *decomposition.* Each has its advantages and disadvantages. Additionally, there are several ways of finding the LU decomposition of a matrix. We will discuss a "cookebook" form of LU decomposition, and another form which takes advantage of the properties of matrix block multiplication. This second form is, initially, slightly more difficult. However, it is actually a simpler form when using computers to perform LU decomposition on large systems. We will look at the errors and costs of the three methods, in terms of the number of computer computations required and the time required to perform the calculation.

## A. SYSTEMS OF LINEAR EQUATIONS

A *linear equation* is an equation which can be written in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where the coefficients $a_i$ and $b$ are real constants and the $x_i$ are variables, or unknowns. The vector $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$ is a *solution* if and only if it satisfies the linear equation.

**Definition IV.1** An equation of the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ is *linear* if the following properties hold:

1. Every variable occurs only in the first power.
2. There are no products of variables.
3. No variables are arguments for radical, exponential, logarithmic or trigonometric functions.

**Example IV.1** Given the following equations, determine whether they are linear or nonlinear. If nonlinear, state why.

a) $x_1x_2 + 3x_2 = 1$

b) $2x_1 + 3x_2 = \sin\pi$

c) $\sin x_1 + x_2 = 3$

d) $\sqrt{3}x_1 + \frac{\pi}{2}x_2 = 0$

e) $x_1 + (x_2)^3 = 5$

**Solution:**

a) Nonlinear, because $x_1x_2$ is a product of two variables.

b) Linear. Remember $\sin\pi$ is a constant.

c) Nonlinear, because $x_1$ is the argument of trigonometric function.

d) Linear.

e) Nonlinear, because $x_2$ occurs to a power higher than 1.

The linear equation $3x_1 - x_2 = 6$, represents a line in $\mathcal{R}^2$ and has as one solution $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T = \begin{bmatrix} 2 & 0 \end{bmatrix}^T$. This solution is not unique. In other words, there are other values for $x_1$ and $x_2$ which satisfy the linear equation. Can you find them? Try $\mathbf{x} = \begin{bmatrix} 3 & 3 \end{bmatrix}^T$.

We can form a *system of linear equations,* which might describe a real world problem, using several linear equations.

$$\begin{array}{ccccc}
a_{1,1}x_1+ & a_{1,2}x_2+ & \cdots+ & a_{1m}x_n = & b_1 \\
a_{2,1}x_1+ & a_{2,2}x_2+ & \cdots+ & a_{2m}x_n = & b_2 \\
. & . & . & . & . \\
. & . & . & . & . \\
a_{n1}x_1+ & a_{n2}x_2+ & \cdots+ & a_{nm}x_n = & b_n
\end{array}$$

By making a slight modification, the above system becomes the *vector equation,*

$$x_1\begin{bmatrix} a_{1,1} \\ a_{2,1} \\ . \\ . \\ a_{n1} \end{bmatrix} + x_2\begin{bmatrix} a_{1,2} \\ a_{2,2} \\ . \\ . \\ a_{n2} \end{bmatrix} + \cdots + x_n\begin{bmatrix} a_{1m} \\ a_{2m} \\ . \\ . \\ a_{nm} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ b_n \end{bmatrix}.$$

When a system of linear equations is written in the vector form, we can more easily discuss the geometric interpretation of adding and subtracting vectors. This form is also helpful when discussing linear combinations of vectors. The matrix equation $\mathbf{Ax} = \mathbf{b}$ is another way of representing a system of linear equations and has its own desirable qualities.

If we make another slight modification to the system of linear equations we get the *matrix equation,*

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2m} \\ . & . & . & . \\ . & . & . & . \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ b_n \end{bmatrix}.$$

This form allows us to perform matrix algebra on the system of linear equations in order to determine the solution. A disadvantage of this form as opposed to the vector

form is that our geometric interpretation of vectors and the idea of linear combinations of vectors seems lost. If we remember to think of columns as vectors, then our interpretations and ideas concerning vectors will also hold true for the columns of A in the matrix equation. Taking this a step further, our interpretations and ideas concerning vectors can be extended to the system of linear equations. We can take this step since the vector equation, matrix equation and the system of linear equations are really different ways of representing the same thing. The different forms help us discuss different ideas.

**Definition IV.2** Given the *system of linear equations* of the form,

$$\begin{array}{ccccc} a_{1,1}x_1+ & a_{1,2}x_2+ & \cdots+ & a_{1m}x_n = & b_1 \\ a_{2,1}x_1+ & a_{2,2}x_2+ & \cdots+ & a_{2m}x_n = & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1}x_1+ & a_{n2}x_2+ & \cdots+ & a_{nm}x_n = & b_n, \end{array}$$

the *matrix of coefficients* is

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix},$$

and the *augmented matrix* is

$$\left[\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1m} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & b_n \end{array}\right].$$

Notice that the coefficients of the system of linear equations form the matrix of coefficients. When we append the constants from the right-hand side of the system of linear equations to the matrix of coefficients, we get the augmented matrix.

Before we demonstrate how to solve systems of linear equations, we must first define some terms which describe our system.

**Definition IV.3** If the system has more variables than equations $(m > n)$, then the system is said to be *underdetermined.* If the system has fewer variables than equations $(m < n)$, then the system is said to be *overdetermined.*

**Example IV.2** Given the following system of linear equations, determine whether the system is underdetermined or overdetermined.

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 3 \\ 2x_1 & + & 2x_2 & + & 2x_3 & = & 1 \end{array}$$

**Solution:** The system is underdetermined, since there are more variables than equations.

**Example IV.3** Given the following system of linear equations, determine whether the system is underdetermined or overdetermined.

$$\begin{array}{rcrcl} x_1 & + & x_2 & = & 1 \\ 2x_1 & + & 2x_2 & = & 2 \\ 3x_1 & + & 3x_2 & = & 3 \\ 4x_1 & + & 4x_2 & = & 4 \\ 5x_1 & + & 5x_2 & = & 5 \\ 6x_1 & + & 6x_2 & = & 6 \end{array}$$

**Solution:** The system is overdetermined, since there are more equations than variables. Regardless of whether a system is overdetermined or underdetermined, a solution is not guaranteed.

**Definition IV.4** If a solution to a system of linear equations exists, then the system is *consistent.* If no solutions exist, then the system is *inconsistent.*

**Example IV.4** Given $0x_1 + 0x_2 = 2$, solve for $\mathbf{x}$.

**Solution:** The equation is inconsistent. Whatever values are chosen for the variables, the left-hand side of the equality is always 0 while the right-hand side is 2.

We have shown what a linear system is and the different ways to represent it. We have also shown that a system may or may not have a solution. Now let's move on to finding a solution, if it exists. First, we start with the *method of substitution.* This method is perhaps the most labor-intensive of all the methods we will show.

**Example IV.5** Given the following system of linear equations, determine a solution by the method of substitution.

$$\begin{array}{rcrcl} 2x_1 & + & 3x_2 & = & 6 \\ x_1 & + & x_2 & = & 2 \end{array}$$

**Solution:** Solve for $x_1$ in the first equation:

$$\begin{aligned} 2x_1 + 3x_2 &= 6 \\ 2x_1 &= 6 - 3x_2 \\ x_1 &= 3 - \frac{3}{2}x_2. \end{aligned}$$

Substitute $x_1 = 3 - \frac{3}{2}x_2$ into the second equation:

$$\begin{aligned} x_1 + x_2 &= 2 \\ 3 - \frac{3}{2}x_2 + x_2 &= 2 \\ -\frac{1}{2}x_2 &= -1 \\ x_2 &= 2. \end{aligned}$$

Substitute $x_2 = 2$ into either of the original equations:

$$\begin{aligned} x_1 + 2 &= 2 \\ x_1 &= 0. \end{aligned}$$

The solution to the system is $\mathbf{x} = \begin{bmatrix} 0 & 2 \end{bmatrix}^T$.

As the numbers of equations and variables increase, the method of substitution can become overwhelming. Imagine writing $x_{10}$ in terms of $x_9, x_8, x_7, x_6, x_5, x_4, x_3, x_2$, and $x_1$. This is obviously not the most efficient method.

## B. GAUSSIAN ELIMINATION

Gaussian elimination is difficult to explain in plain English, so we will use examples to demonstrate the process. First lets look at a system of linear equations and try eliminating variables.

**Example IV.6** Given the following system of linear equations, compute the solution **x**.

$$\begin{array}{rcrcrcr} 2x_1 & + & 4x_2 & - & 2x_3 & = & -4 \\ 3x_1 & + & x_2 & - & 2x_3 & = & -5 \\ 9x_1 & + & 6x_2 & - & 6x_3 & = & 3 \end{array}$$

**Solution:** Start by dividing the first equation by 2 and the last equation by 3. This gives

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & - & 1x_3 & = & -2 \\ 3x_1 & + & x_2 & - & 2x_3 & = & -5 \\ 3x_1 & + & 2x_2 & - & 2x_3 & = & 1. \end{array}$$

Eliminate $x_1$ by multiplying the first equation by $-3$ and adding it to the other equations. This gives

$$\begin{array}{rcrcrcr} 1x_1 & + & 2x_2 & - & x_3 & = & -2 \\ & - & 5x_2 & + & x_3 & = & 1 \\ & - & 4x_2 & + & x_3 & = & 7. \end{array}$$

Eliminate $x_2$ by multiplying the second equation by $-\frac{4}{5}$ and adding it to the last equation, which gives us a system easily solved:

$$\begin{array}{rcrcrcr} 1x_1 & + & 2x_2 & - & x_3 & = & -2 \\ & - & 5x_2 & + & x_3 & = & 1 \\ & & & + & \frac{1}{5}x_3 & = & \frac{31}{5}. \end{array}$$

We can perform this process of eliminating variables on an augmented matrix using elementary row operations. This process is known as *Gaussian elimination.* The following are elementary row operations.

a) Scaling: Multiplication of a row by a scalar.

b) Row interchange: Interchange of two rows.

c) Row replacement: Addition of a scalar multiple of one row to another row.

The previous system of linear equations can be written as

$$\left[\begin{array}{rrr|r} 1 & 2 & -1 & -2 \\ 0 & -5 & 1 & 1 \\ 0 & 0 & \frac{1}{5} & \frac{31}{5} \end{array}\right],$$

and is said to be in *row echelon form.* An augmented matrix is in row echelon form when all of the following are true:

a) Any row or rows consisting entirely of zeros occurs as the last row or rows.

b) The leading non-zero entry of any row, known as the *pivot,* is to the left of all pivots in subsequent rows.

c) All entries below pivots are zero.

**Example IV.7** Given $A = \left[\begin{array}{rrr|r} 6 & 4 & -2 & -4 \\ 3 & 1 & -2 & -5 \\ 9 & 6 & 3 & 6 \end{array}\right]$, use elementary row operations to reduce it to row echelon form.

**Solution:** Replace row2 by $-\frac{3}{6}$row1 + row2 and replace row3 by $-\frac{9}{6}$row1 + row3. This gives

$$\left[\begin{array}{rrr|r} 6 & 4 & -2 & -4 \\ 0 & -1 & -1 & -3 \\ 0 & 0 & 6 & 12 \end{array}\right].$$

The augmented matrix is now in row echelon form and is itself a system of linear equations that can be expressed as

$$\begin{array}{rcrcrcr} 6x_1 & + & 4x_2 & - & 2x_3 & = & -4 \\ & - & x_2 & - & x_3 & = & -3 \\ & & & & 6x_3 & = & 12. \end{array}$$

**Definition IV.5** If two augmented matrices have the same solution set, then they are said to be *equivalent.*

**Theorem IV.1** *If any elementary row operation is performed on a augmented matrix* $A \in \mathcal{R}^{n \times m}$*, then the resulting augmented matrix has the same solution set as the original augmented matrix.*

Using this theorem, we can verify that our row echelon form matrix from the previous example is equivalent to the original matrix. In order to determine the solution set, we will solve for the variable in the last row and then substitute it into the row above to solve for the next variable and repeat. This process is called *back-substitution* for obvious reasons.

**Example IV.8** Given $A = \left[\begin{array}{ccc|c} 6 & 4 & -2 & -4 \\ 0 & -1 & -1 & -3 \\ 0 & 0 & 6 & 12 \end{array}\right]$ from the previous example, use back-substitution to compute the solution set.

**Solution:** Writing this augmented matrix as a system of linear equations yields

$$\begin{array}{rcrcrcr} 6x_1 & + & 4x_2 & - & 2x_3 & = & -4 \\ & - & x_2 & - & x_3 & = & -3 \\ & & & & 6x_3 & = & 12. \end{array}$$

Starting with the last equation, solve for $x_3$, then use repeated back-substitution to solve for $x_2$ and $x_1$. Therefore, from the equation $6x_3 = 12$, we have $x_3 = 2$. Back substitute $x_3 = 2$ into the second equation:

$$\begin{aligned} -x_2 - x_3 &= -3 \\ -x_2 - 2 &= -3 \\ -x_2 &= -3 + 2 \\ x_2 &= 1. \end{aligned}$$

Back substitute $x_2 = 1$ and $x_3 = 2$ into the first equation:

$$\begin{aligned} 6x_1 + 4x_2 - 2x_3 &= -4 \\ 6x_1 + 4(1) - 2(2) &= -4 \\ 6x_1 &= -4 - 4 + 4 \\ x_1 &= -\frac{2}{3}. \end{aligned}$$

Therefore, the solution is $\mathbf{x} = \begin{bmatrix} -\frac{2}{3} & 1 & 2 \end{bmatrix}^T$. Notice that this solution is unique. Also, since a solution does exist, the system of linear equations is consistent. If the solution is correct, each equation should be satisfied simultaneously as follows:

$$\begin{array}{rcrcrcrcrcrcrcr}
6x_1 & + & 4x_2 & - & 2x_3 & = & 6(-\frac{2}{3}) & + & 4(1) & - & 2(2) & = & -4 \\
3x_1 & + & x_2 & - & 2x_3 & = & 3(-\frac{2}{3}) & + & 1 & - & 2(2) & = & -5 \\
9x_1 & + & 6x_2 & + & 3x_3 & = & 9(-\frac{2}{3}) & + & 6(1) & + & 3(2) & = & 6.
\end{array}$$

Therefore, $\mathbf{x}$ is the solution to the original system.

**Example IV.9** Given $A = \left[\begin{array}{cc|c} i & -2 & 3 \\ 1-i & 2-3i & 0 \end{array}\right]$, use Gaussian elimination to reduce it to row echelon form.

**Solution:** Multiply row1 by $-i$ and row2 by $1+i$. This makes the elements in the first column of A real. The result is

$$\left[\begin{array}{cc|c} 1 & 2i & -3i \\ 2 & 5-i & 0 \end{array}\right].$$

Using Gaussian elimination, replace row2 with $-2$row1 + row2, which yields the row echelon form

$$\left[\begin{array}{cc|c} 1 & 2i & -3i \\ 0 & 5-5i & 6i \end{array}\right].$$

**Example IV.10** Given $A = \left[\begin{array}{cc|c} 1 & 2i & -3i \\ 0 & 5-5i & 6i \end{array}\right]$, use back-substitution to compute the solution set.

**Solution:** Multiply row2 by $5+5i$ to make the element in the second column, second row real. This gives

$$A = \left[\begin{array}{cc|c} 1 & 2i & -3i \\ 0 & 50 & -30+30i \end{array}\right].$$

Therefore, from the second row, we have $50x_2 = -30+30i$, which yields $x_2 = \frac{-3+3i}{5}$.

Back substitute $x_2 = \frac{-3+3i}{5}$ into the second equation:

$$\begin{aligned} x_1 + 2ix_2 &= -3i \\ x_1 + 2i\left(\frac{-3+3i}{5}\right) &= -3i \\ x_1 &= \frac{6-9i}{5}. \end{aligned}$$

Therefore, the solution is $\mathbf{x} = \begin{bmatrix} \frac{6-9i}{5} & \frac{-3+3i}{5} \end{bmatrix}^T$.

By applying a few more elementary row operations to an augmented matrix, we can compute the solution directly instead of using back-substitution.

**Example IV.11** Given A $= \left[\begin{array}{ccc|c} 6 & 4 & -2 & -4 \\ 0 & -1 & -1 & -3 \\ 0 & 0 & 6 & 12 \end{array}\right]$, compute the solution set.

**Solution:** Divide row3 by 6 to obtain

$$\left[\begin{array}{ccc|c} 6 & 4 & -2 & -4 \\ 0 & -1 & -1 & -3 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

Replace row2 by row3 + row2 and replace row1 by 2row3 + row1; this yields

$$\left[\begin{array}{ccc|c} 6 & 4 & 0 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

Replace row1 by 4row2 + row1 then replace row2 by −row2 to get

$$\left[\begin{array}{ccc|c} 6 & 0 & 0 & -4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

Replace row1 by $\frac{1}{6}$row1 to get

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{2}{3} \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

Solving for the variables directly yields $x_1 = -\frac{2}{3}, x_2 = 1$, and $x_3 = 2$, or $\mathbf{x} = \begin{bmatrix} -\frac{2}{3} & 1 & 2 \end{bmatrix}^T$. The elimination process that we performed is called *Gauss-Jordan elimination.* It puts the augmented matrix into *reduced row echelon form* so that the solution can be found by directly solving for each variable vice using back-substitution. An augmented matrix is in reduced row echelon form when all of the following are true:

a) A is in row-echelon form.

b) All pivots are 1.

c) All entries above the pivots are zero.

We have solved systems of linear equations by method of substitution, and their associated augmented matrices by Gaussian and Gauss-Jordan elimination. But are these solutions valid in the vector and matrix equation forms? Lets look at two examples using the following system of linear equations, that shows the solution set to the system is also valid for the equivalent vector and matrix equations.

$$\begin{array}{rcrcrcr} 6x_1 & + & 4x_2 & - & 2x_3 & = & -4 \\ 3x_1 & + & x_2 & - & 2x_3 & = & -5 \\ 9x_1 & + & 6x_2 & + & 3x_3 & = & 6 \end{array}$$

As shown previously, the solution to this system is $\mathbf{x} = \begin{bmatrix} -\frac{2}{3} & 1 & 2 \end{bmatrix}^T$.

**Example IV.12** Given the vector equation $x_1 \begin{bmatrix} 6 \\ 3 \\ 9 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 1 \\ 6 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ -5 \\ 6 \end{bmatrix}$, determine whether $\mathbf{x} = \begin{bmatrix} -\frac{2}{3} & 1 & 2 \end{bmatrix}^T$ is a solution.

**Solution:** Plugging $\mathbf{x}$ into the vector equation yields

$$-\frac{2}{3} \begin{bmatrix} 6 \\ 3 \\ 9 \end{bmatrix} + 1 \begin{bmatrix} 4 \\ 1 \\ 6 \end{bmatrix} + 2 \begin{bmatrix} -2 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ -5 \\ 6 \end{bmatrix}.$$

We see that $\mathbf{x}$ is a solution.

**Example IV.13** Given the matrix equation $\begin{bmatrix} 6 & 4 & -2 \\ 3 & 1 & -2 \\ 9 & 6 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ -5 \\ 6 \end{bmatrix}$, determine whether $\mathbf{x} = \begin{bmatrix} -\frac{2}{3} & 1 & 2 \end{bmatrix}^T$ is a solution.

**Solution:** Plugging $\mathbf{x}$ into the matrix equation yields

$$\begin{bmatrix} 6 & 4 & -2 \\ 3 & 1 & -2 \\ 9 & 6 & 3 \end{bmatrix} \begin{bmatrix} -\frac{2}{3} \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ -5 \\ 6 \end{bmatrix}.$$

We see that $\mathbf{x}$ is a solution.

Now that we can solve a system of linear equations, we define some terms which apply to a reduced augmented matrix.

**Definition IV.6** Given an augmented matrix in row echelon or reduced row echelon form, the variables which correspond to pivots are called *pivot variables.* The remaining variables are called *free variables,* and can take on real parametric values.

**Definition IV.7** The set of all vectors $\mathbf{x}$ that simultaneously satisfies all equations in a system of linear equations is said to be the *solution set* of the system.

**Example IV.14** Given the following augmented matrix in row echelon form, determine the pivot variables, free variables and the solution set.

$$A = \left[\begin{array}{cccccc|c} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

**Solution:** The variables $x_1, x_3$, and $x_6$ are pivot variables, since they occupy the leading non-zero entry of rows 1, 2, and 3. The remaining variables $x_2, x_4$, and $x_5$ are free variables. In order to perform back-substitution on the augmented matrix, we must first set the free variables $x_2, x_4$, and $x_5$ equal to some parametric values. We will use $r, s$, and $t \in \mathcal{R}$, respectively. Then, we can use back-substitution to solve for the pivot variables in terms of the free variables. The last row, because it

consists entirely of zeros, adds no information. Therefore, starting with row three, we directly solve for $x_6$, which yields $x_6 = \frac{1}{3}$. Substitute $x_6 = \frac{1}{3}$ into row2, and make the parametric substitutions $x_4 = s$ and $x_5 = t$, which yields

$$\begin{aligned} x_3 + 2s + 0t + 3(\frac{1}{3}) &= 4 \\ x_3 &= 4 - 2s - 1 = 3 - 2s. \end{aligned}$$

Substitute $x_6 = \frac{1}{3}, x_5 = t, x_4 = s, x_3 = 3 - 2s$, and $x_2 = r$ into row1 and solve for $x_1$ :

$$\begin{aligned} x_1 + 3r - 2(3 - 2s) + 0s + 2t + 0\frac{1}{3} &= 0 \\ x_1 &= -3r + 6 - 4s - 2t = 6 - 3r - 4s - 2t. \end{aligned}$$

The solution set is then $\mathbf{x} = \begin{bmatrix} 6 - 3r - 4s - 2t & r & 3 - 2s & s & t & \frac{1}{3} \end{bmatrix}^T$, where $r, s,$ and $t \in \mathcal{R}$. By assigning the parameters $r, s,$ and $t$ different values, we can find infinitely many solutions to the system. Note that because a solution exists, the system is consistent.

**Example IV.15** Given the following augmented matrix in row echelon form, compute the solution set.

$$\mathrm{A} = \left[\begin{array}{cccc|c} 1 & 3 & -2 & 2 & 0 \\ 0 & 5 & 1 & 0 & 4 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 6 & 2 \end{array}\right].$$

**Solution:** $x_1, x_2, x_3$, and $x_4$ are pivot variables. This system has no free variables. Using back-substitution, we find that $\mathbf{x} = \begin{bmatrix} \frac{76}{15} & \frac{4}{5} & -1 & \frac{1}{3} \end{bmatrix}^T$. Because a solution exists, the system is consistent. Also, there are no parametric values, therefore the solution is unique.

**Example IV.16** Given the following augmented matrix in row echelon form, determine the solution set.

$$\mathrm{A} = \left[\begin{array}{cccccc|c} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right].$$

**Solution:** $x_1, x_3$, and $x_6$ are pivot variables and $x_2, x_4$, and $x_5$ are free variables. But, in the last row we have $0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 + 0x_6 = 2$, which means that no solution exists, and therefore the system is inconsistent.

In previous examples we saw systems with no free variables, which led to a unique solution. We also saw systems with free variables, which led to either an infinite number of solutions or no solutions.

## C. EXISTENCE AND UNIQUENESS OF A SOLUTION

In the previous section we reduced the augmented matrices into their row echelon or reduced row echelon form. Then we solved the system using back-substitution. However, a solution did not always exist. We can determine whether a solution exists or not by looking at the last rows of a row echelon matrix or a reduced row echelon matrix. If a solution does indeed exist, we can also determine whether it is unique or one of infinitely many. Recall that if a system of linear equations has any equations of the form

$$0x_1 + 0x_2 + \cdots + 0x_n = b, \ b \neq 0,$$

then the system of linear equations is inconsistent. However, the above equation may not appear in the augmented matrix until after Gaussian elimination is complete and the augmented matrix is in row echelon or reduced row echelon form.

**Example IV.17** Given $\mathrm{A} = \left[\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{4}{6} \\ -2 & 1 & 0 & \frac{4}{3} \\ 1 & 1 & 0 & \frac{1}{3} \end{array}\right]$, determine the solution set.

**Solution:** After performing Gaussian elimination, the row echelon form is

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{4}{6} \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right].$$

In the last row we have the equation $0x_1 + 0x_2 + 0x_3 = 2$. No matter what values we choose for $x_1, x_2$, and $x_3$, the equation cannot be satisfied. Therefore, no solution exists to the system.

**Example IV.18** Given A $= \left[\begin{array}{rrr|r} 7 & 0 & -3 & -6 \\ 7 & 9 & 3 & 3 \\ -7 & 0 & 0 & 2 \end{array}\right]$, determine the solution set.

**Solution:** After performing Gaussian elimination, the row echelon form is

$$\left[\begin{array}{rrr|r} 7 & 0 & -3 & -6 \\ 0 & 9 & 0 & -3 \\ 0 & 0 & 3 & 8 \end{array}\right].$$

This augmented matrix does not contain any rows of the form $0x_1 + 0x_2 + \cdots + 0x_n = b$, $b \neq 0$. Therefore, a solution exists, and we can solve for each pivot variable using back-substitution. Notice that the solution $\mathbf{x} = \begin{bmatrix} \frac{2}{7} & -\frac{1}{3} & \frac{8}{3} \end{bmatrix}^T$ is unique and that the number of pivot variables is equal to the number of equations.

**Example IV.19** Given A $= \left[\begin{array}{rrr|r} 9 & 9 & -6 & -2 \\ 0 & 3 & -6 & 5 \\ 9 & 12 & -12 & 3 \end{array}\right]$, determine the solution set.

**Solution:** After performing Gaussian elimination, the row echelon form is

$$\left[\begin{array}{rrr|r} 9 & 6 & 0 & -7 \\ 0 & 3 & -6 & 5 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

This augmented matrix does not contain any rows of the form $0x_1 + 0x_2 + \cdots + 0x_n = b$, $b \neq 0$. Therefore, a solution exists. Notice that there are two pivot variables and one free variable. If we let $x_3$ be the parametric value $t$, then back substitution yields the solution set $\mathbf{x} = \begin{bmatrix} -\frac{17+12t}{9} & \frac{5+6t}{3} & t \end{bmatrix}^T$, where $t \in \mathcal{R}$. Notice that there are fewer pivot variables than equations.

Let's summarize what we have just learned from these examples. Given the row echelon or reduced row echelon form of the augmented matrix, if there is an

inconsistent row, then the system has no solution. If the number of pivot variables is equal to the number of equations, then the solution is unique. If there are free variables, then there are infinitely many solutions.

If the right-hand column of an augmented matrix is all zeros, then the system is said to be *homogeneous.* If the right-hand side is non-zero, then the system is said to be *non-homogeneous.* The zero vector, or *trivial solution*, is always a solution to the homogeneous system. A non-trivial solution may or may not exist.

**Example IV.20** Given the following homogeneous, augmented matrix in row echelon form, determine the solution set.

$$A = \left[\begin{array}{cccccc|c} -1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

**Solution:** As stated previously, A has the trivial solution $\mathbf{x} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$. There are three pivot variables, and three free variables which means we can find a non-trivial solution. As before, we set the free variables equal to parameters $r, s$, and $t \in \mathcal{R}$. Then we use back-substitution to solve for the pivot variables, which yields $\mathbf{x} = \begin{bmatrix} 3r - 4s & r & 2s + t & s & t & 0 \end{bmatrix}^T$.

## D. INVERSE OF A MATRIX

Repeated Gaussian elimination on a large matrix is inefficient in terms of the number of computer computations. Remember that our augmented matrix comes from some real-world problem with potentially hundreds or thousands of variables. Therefore, repeated Gaussian elimination is not the method of choice for solving systems of linear equations. Let's build a better mousetrap using the inverse of a matrix.

So far we have looked at addition, subtraction and multiplication of matrices. The obvious question now, is what about division? After all, we have developed all

the other basic algebraic operations, and division of matrices is certainly one of them. Unfortunately, there is no natural way to talk about division of matrices. To get a better understanding, we need to go back to some very basic notions about algebraic operations, and how these basic operations work.

Addition and subtraction are natural concepts. We see them in our everyday life whenever we count. Multiplication is artificial, it is really just a quick way to perform repeated addition. Division is also artificial and comes from the need to solve linear equations like $ax = b$ given that you know the scalars $a$ and $b$. In fact, you do not have to know general division to solve this problem, provided you know how to compute $a^{-1}$. The reciprocal of $a$, also known as the multiplicative inverse, enables us to solve linear equations using multiplication; $x = a^{-1}b$.

**Example IV.21** Given $2x = 6$, solve for $x$.

**Solution:** Multiply both sides of the equation by the multiplicative inverse $\frac{1}{2}$,which yields $x = 3$.

**Definition IV.8** Let $A, B \in \mathcal{R}^{n \times n}$. If $AB = BA = I$, we say that B is the *multiplicative inverse* of A, denoted $B = A^{-1}$.

Unfortunately, there are some difficulties. First of all, since we want $A^{-1}A = AA^{-1}$, it is clear that only square matrices can have inverses.

**Definition IV.9** A matrix is *nonsingular,* or *invertible,* if it has an inverse, and is *singular* if it does not have an inverse.

At this point in the text, it is only important to understand what an inverse is, vice how to find the inverse of a matrix. Later we will learn how to find the inverse of a matrix if it exists.

We can solve the matrix equation $Ax = \mathbf{b}$ by left-multiplying the equation by $A^{-1}$ :

$$\begin{aligned}A^{-1}Ax &= A^{-1}\mathbf{b}\\ Ix &= A^{-1}\mathbf{b}\\ x &= A^{-1}\mathbf{b}.\end{aligned}$$

Now we need to find $A^{-1}$. We've already shown how to find the inverse of a scalar. Since only square matrices have inverses, the next obvious step is to find the inverse of a $2 \times 2$ matrix. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then we can use the formula:

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \text{ provided } ad - bc \neq 0.$$

Notice that if $ad - bc = 0$, then $A^{-1}$ does not exist.

**Example IV.22** Given $A= \begin{bmatrix} 2 & -1 \\ 5 & 3 \end{bmatrix}$, compute $A^{-1}$.

**Solution:**

$$A^{-1} = \frac{1}{6+5}\begin{bmatrix} 3 & 1 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{11} & \frac{1}{11} \\ -\frac{5}{11} & \frac{2}{11} \end{bmatrix}.$$

This is easily verified:

$$\begin{aligned}A^{-1}A &= \begin{bmatrix} \frac{3}{11} & \frac{1}{11} \\ -\frac{5}{11} & \frac{2}{11} \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 5 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \text{ and}\\ AA^{-1} &= \begin{bmatrix} 2 & -1 \\ 5 & 3 \end{bmatrix}\begin{bmatrix} \frac{3}{11} & \frac{1}{11} \\ -\frac{5}{11} & \frac{2}{11} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.\end{aligned}$$

For matrices larger than $2 \times 2$, we will use Gauss-Jordan elimination. Let's show why this will work using elementary row replacement matrices to reduce an invertible matrix A to the identity matrix:

$$E_k \cdots E_3 E_2 E_1 A = I$$

(It can be shown that this is possible whenever $A^{-1}$exists.)

Left multiplying the matrix equation by $E_k \cdots E_3 E_2 E_1$, we have

$$\begin{aligned} E_k \cdots E_3 E_2 E_1 A\mathbf{x} &= E_k \cdots E_3 E_2 E_1 \mathbf{b} \\ I\mathbf{x} &= E_k \cdots E_3 E_2 E_1 \mathbf{b} \\ \mathbf{x} &= E_k \cdots E_3 E_2 E_1 \mathbf{b}. \end{aligned}$$

Solving $A\mathbf{x} = \mathbf{b}$ using the inverse gives the matrix equation $\mathbf{x} = A^{-1}\mathbf{b}$. Substituting $\mathbf{x} = E_k \cdots E_3 E_2 E_1 \mathbf{b}$ into the matrix equation gives $E_k \cdots E_3 E_2 E_1 \mathbf{b} = A^{-1}\mathbf{b}$. Therefore, assuming $\mathbf{b} \neq \mathbf{0}$, $A^{-1} = E_k \cdots E_3 E_2 E_1$. Elementary matrices are cumbersome, so instead we will use Gauss-Jordan elimination to find the inverse of A. This is done as follows: Place the identity matrix to the right of A, to form a new matrix

$$\left[\begin{array}{cccc|cccc} a_{1,1} & a_{1,2} & \cdots & a_{1m} & 1 & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & \cdots & a_{2m} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & 0 & 0 & \cdots & 1 \end{array}\right].$$

Now perform Gauss-Jordan elimination on the entire matrix. The product of the elementary matrices is stored in the right-hand portion of the new matrix, $[I|E_k \cdots E_3 E_2 E_1]$. By the preceeding remarks, the matrix on the right is $A^{-1}$.

**Example IV.23** Given $A = \begin{bmatrix} 4 & -2 & 6 \\ 2 & 0 & -4 \\ 2 & 0 & 2 \end{bmatrix}$, compute $A^{-1}$.

**Solution:** Form a new matrix using A and I, and then perform Gauss-Jordan elimination:

$$\left[\begin{array}{ccc|ccc} 4 & -2 & 6 & 1 & 0 & 0 \\ 2 & 0 & -4 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 & 0 & 1 \end{array}\right].$$

Replace row2 by $-\frac{1}{2}$row1 + row2 and replace row3 by $-\frac{1}{2}$row1 + row3, which gives

$$\left[\begin{array}{ccc|ccc} 4 & -2 & 6 & 1 & 0 & 0 \\ 0 & 1 & -7 & -\frac{1}{2} & 1 & 0 \\ 0 & 1 & -1 & -\frac{1}{2} & 0 & 1 \end{array}\right].$$

Replace row3 by $-$row2 + row3, which gives

$$\left[\begin{array}{ccc|ccc} 4 & -2 & 6 & 1 & 0 & 0 \\ 0 & 1 & -7 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 6 & 0 & -1 & 1 \end{array}\right].$$

Replace row3 by $\frac{1}{6}$row3, replace row2 by 7row3 + row2 and replace row1 by $-6$row3 + row1, which gives

$$\left[\begin{array}{ccc|ccc} 4 & -2 & 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & -\frac{1}{2} & -\frac{1}{6} & \frac{7}{6} \\ 0 & 0 & 1 & 0 & -\frac{1}{6} & \frac{1}{6} \end{array}\right].$$

Replace row1 by 2row2 + row1 and replace row1 by $\frac{1}{4}$row1, which gives

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{3} \\ 0 & 1 & 0 & -\frac{1}{2} & -\frac{1}{6} & \frac{7}{6} \\ 0 & 0 & 1 & 0 & -\frac{1}{6} & \frac{1}{6} \end{array}\right].$$

The inverse matrix is

$$A^{-1} = \begin{bmatrix} 0 & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{2} & \frac{-1}{6} & \frac{7}{6} \\ 0 & \frac{-1}{6} & \frac{1}{6} \end{bmatrix}.$$

**Example IV.24** Given the following matrix, compute $A^{-1}$.

$$A = \begin{bmatrix} 3 & 0 & -3 & 9 \\ 2 & 4 & 0 & -1 \\ 0 & -1 & 6 & 4 \\ 7 & 8 & -3 & 6 \end{bmatrix}$$

**Solution:** In order to simplify some of the calculations, ensure that each pivot is 1.

$$\begin{bmatrix} 3 & 0 & -3 & 9 & 1 & 0 & 0 & 0 \\ 2 & 4 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 6 & 4 & 0 & 0 & 1 & 0 \\ 7 & 8 & -3 & 6 & 0 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & 3 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 4 & 2 & -7 & \frac{-2}{3} & 1 & 0 & 0 \\ 0 & -1 & 6 & 4 & 0 & 0 & 1 & 0 \\ 0 & 8 & 4 & -15 & \frac{-7}{3} & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & -1 & 3 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{-7}{4} & \frac{-1}{6} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{13}{2} & \frac{9}{4} & \frac{-1}{6} & \frac{1}{4} & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & -2 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & 3 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{-7}{4} & \frac{-1}{6} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 1 & \frac{18}{52} & \frac{-2}{78} & \frac{2}{52} & \frac{2}{13} & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 & 0 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & -1 & 0 & \frac{-8}{3} & -6 & 0 & -3 \\ 0 & 1 & \frac{1}{2} & 0 & \frac{19}{12} & \frac{15}{4} & 0 & \frac{-7}{4} \\ 0 & 0 & 1 & 0 & \frac{-58}{156} & \frac{-34}{52} & \frac{2}{13} & \frac{18}{52} \\ 0 & 0 & 0 & 1 & 1 & 2 & 0 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & 0 & \frac{-8}{3} & -6 & 0 & 3 \\ 0 & 1 & 0 & 0 & \frac{276}{156} & \frac{212}{52} & \frac{-2}{26} & \frac{-100}{52} \\ 0 & 0 & 1 & 0 & \frac{-58}{156} & \frac{-34}{52} & \frac{2}{13} & \frac{18}{52} \\ 0 & 0 & 0 & 1 & 1 & 2 & 0 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \frac{-474}{156} & \frac{-346}{52} & \frac{2}{13} & \frac{174}{52} \\ 0 & 1 & 0 & 0 & \frac{276}{156} & \frac{212}{52} & \frac{-2}{26} & \frac{-100}{52} \\ 0 & 0 & 1 & 0 & \frac{-58}{156} & \frac{-34}{52} & \frac{2}{13} & \frac{18}{52} \\ 0 & 0 & 0 & 1 & 1 & 2 & 0 & -1 \end{bmatrix}$$

We have combined a few of the steps to save space. During the elimination process we can perform row replacements and scalar multiplication only. Row exchanges require permutation matrices. In order to maintain simplicity we will not perform row exchanges.

## E. TRIANGULAR SYSTEMS

Triangular systems have a very useful form. We have already shown how this form can be utilized to solve systems of linear equations. That is, every time we reduce our augmented matrix into its row echelon or reduced row echelon form, we have reduced it to a triangular matrix. In the next section, triangular matrices will be used often when performing back-substitution and forward elimination.

**Definition IV.10** In an *upper triangular* matrix, all of the entries below the main diagonal are zero.

**Example IV.25** A is an upper triangular matrix:

$$A = \begin{bmatrix} 2 & -1 & 0 & 5 & 3 & 1 \\ 0 & 7 & 3 & 0 & 8 & -6 \\ 0 & 0 & 0 & -1 & 6 & 4 \\ 0 & 0 & 0 & 0 & 3 & 2 \end{bmatrix}.$$

**Definition IV.11** In a *lower triangular* matrix, all of the entries above the main diagonal are zero.

**Example IV.26** A is a lower triangular matrix:

$$A = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 5 & 2 & 0 & 0 & 0 \\ -3 & 0 & -4 & 0 & 0 \\ 2 & -5 & 9 & 0 & 0 \end{array}\right].$$

## F. LU DECOMPOSITION

LU decomposition is another method used to solve $A\mathbf{x} = \mathbf{b}$. In real-world problems, $\mathbf{b}$ may change. If Gaussian elimination is used, it would need to be repeated for each new $\mathbf{b}$. However, LU decomposition allows us to change $\mathbf{b}$ without repeating Gaussian elimination. First we will show how to solve $LU\mathbf{x} = \mathbf{b}$. Then we will discuss how to find L and U.

Let $A = LU$, where L is a lower triangular matrix and U is an upper triangular matrix. Then the matrix equation becomes $LU\mathbf{x} = \mathbf{b}$. In this form we can find

a solution by performing a simple forward elimination and then an equally simple back-substitution. First we let $U\mathbf{x} = \mathbf{z}$, which gives

$$A\mathbf{x} = LU\mathbf{x} = L\mathbf{z} = \mathbf{b}.$$

Since L is lower triangular, we can use forward elimination to solve for $\mathbf{z}$, and then perform back-substitution on $U\mathbf{x} = \mathbf{z}$. If $\mathbf{b}$ changes, we will have to repeat the forward elimination and back-substitution using the LU we have already found. If $\mathbf{b}$ changes for any reason in the augmented matrix we will have to repeat Gaussian elimination and back-substitution. After a few homework exercises you will agree that forward elimination is less costly than Gaussian elimination in terms of computations and time. Now all we have to do is find a L and a U which satisfy A = LU. When we performed Gaussian elimination on the augmented matrix, the result was an upper triangular system. Then we used back-substitution to solve the system. Here, we will use Gaussian elimination on A, instead of on the augmented matrix.

**Example IV.27** Given $A = \left[\begin{array}{cc|c} 2 & -3 & 1 \\ 1 & 0 & 2 \end{array}\right]$, use both Gaussian elimination and LU decomposition to compute the solution set.

**Solution:** Beginning with Gaussian elimination, replace row2 by $-\frac{1}{2}$row1 + row2, which gives

$$\left[\begin{array}{cc|c} 2 & -3 & 1 \\ 0 & \frac{3}{2} & \frac{3}{2} \end{array}\right].$$

Starting with the last equation, solve for $x_2$, then use back-substitution to solve for $x_1$. From the equation $\frac{3}{2}x_2 = \frac{3}{2}$, we get $x_2 = 1$. We now substitute $x_2 = 1$ into the first equation:

$$\begin{aligned} 2x_1 - 3(1) &= 1 \\ 2x_1 &= 4 \\ x_1 &= 2. \end{aligned}$$

Therefore, $\mathbf{x} = \begin{bmatrix} 2 & 1 \end{bmatrix}^T$.

Now we begin the LU decomposition by performing Gaussian elimination on A to get an upper triangular matrix:

$$\mathrm{U} = \begin{bmatrix} 2 & -3 \\ 0 & \frac{3}{2} \end{bmatrix}.$$

We now need a lower triangular matrix so that LU = A, so let $\mathrm{L} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}$. (It is easy to verify that LU = A.)

Use forward-elimination on the augmented matrix to solve for $\mathbf{z}$ in $\mathrm{L}\mathbf{z} = \mathbf{b}$,

$$\left[\begin{array}{cc|c} 1 & 0 & 1 \\ \frac{1}{2} & 1 & 2 \end{array}\right].$$

Starting with the first row, solve for $x_1$, then use forward-elimination to solve for $x_2$ in the second row. Therefore, from the first row, $x_1 = 1$. Forward-eliminate $x_1$ by substituting $x_1 = 1$ into the second row:

$$\begin{aligned} \frac{1}{2}(1) + x_2 &= 2 \\ x_2 &= \frac{3}{2}. \end{aligned}$$

Therefore, $\mathbf{x} = \begin{bmatrix} 1 & \frac{3}{2} \end{bmatrix}^T$.

Use back-substitution on the augmented matrix to solve for $\mathbf{x}$ in $\mathrm{U}\mathbf{x} = \mathbf{z}$,

$$\left[\begin{array}{cc|c} 2 & -3 & 1 \\ 0 & \frac{3}{2} & \frac{3}{2} \end{array}\right].$$

Starting with the second equation, solve for $x_2$, then use back-substitution to solve for $x_1$. Therefore, from the equation $\frac{3}{2}x_2 = \frac{3}{2}$, we have $x_2 = 1$. Back-substitute $x_2 = 1$ into the first equation:

$$
\begin{aligned}
2x_1 - 3(1) &= 1 \\
x_1 &= 2.
\end{aligned}
$$

Therefore, $\mathbf{x} = \begin{bmatrix} 1 & \frac{3}{2} \end{bmatrix}^T$. Now let's verify that the L we chose satisfies LU = A.

$$
\mathrm{LU} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ 0 & \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ 1 & 0 \end{bmatrix} = \mathrm{A}.
$$

**Example IV.28** Given $\mathrm{A} = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 4 & 4 \\ -1 & 2 & 5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix}^T$, $\mathrm{L} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$, and $\mathrm{U} = \begin{bmatrix} 1 & 6 & 4 \\ 0 & -8 & -4 \\ 0 & 0 & 5 \end{bmatrix}$, compute the solution set to $\mathrm{A}\mathbf{x} = \mathbf{b}$ using the LU decomposition A = LU.

**Solution:** Solve for $\mathbf{z}$ in $\mathrm{L}\mathbf{z} = \mathbf{b}$, which gives $\mathbf{z} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$.
Solve for $\mathbf{x}$ in $\mathrm{U}\mathbf{x} = \mathbf{z}$, which gives $\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$.
Verify the solution:

$$
\mathrm{A}\mathbf{x} = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 4 & 4 \\ -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \mathbf{b}.
$$

Therefore, our solution is correct.

We get U for free by performing Gaussian elimination on A. But we still need to find an L. Using elementary matrices, we can find a method for building the matrix L. We use elementary matrices to perform Gaussian elimination on A in order to find a $\mathrm{U} = \mathrm{E}_k \cdots \mathrm{E}_3\mathrm{E}_2\mathrm{E}_1\mathrm{A}$. We then left-multiply $\mathrm{A}\mathbf{x} = \mathbf{b}$ by $\mathrm{E}_k \cdots \mathrm{E}_3\mathrm{E}_2\mathrm{E}_1$. This yields

$$
\begin{aligned}
\mathrm{E}_k \cdots \mathrm{E}_3\mathrm{E}_2\mathrm{E}_1\mathrm{A}\mathbf{x} &= \mathrm{E}_k \cdots \mathrm{E}_3\mathrm{E}_2\mathrm{E}_1\mathbf{b} \\
\mathrm{U}\mathbf{x} &= \mathrm{E}_k \cdots \mathrm{E}_3\mathrm{E}_2\mathrm{E}_1\mathbf{b}.
\end{aligned}
$$

Using inverses, we solve for $\mathbf{b}$:

$$[E_k \cdots E_3 E_2 E_1]^{-1} U\mathbf{x} = \mathbf{b}.$$

This last equation should look familiar. It is simply $LU\mathbf{x} = \mathbf{b}$. So let $L = [E_k \cdots E_3 E_2 E_1]^{-1}$. If we use the same principle that we used in finding the inverse of a matrix, then we can find the product of $E_k \cdots E_3 E_2 E_1$. But remember, we will perform Gaussian elimination and not Gauss-Jordan elimination.

**Example IV.29** Given $A = \begin{bmatrix} 1 & 2 & 3 \\ -3 & 2 & 1 \\ 2 & -2 & 1 \end{bmatrix}$, compute LU such that $A = LU$.

**Solution:** Form a new matrix using A and I and then perform Gaussian elimination:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ -3 & 2 & 1 & 0 & 1 & 0 \\ 2 & -2 & 1 & 0 & 0 & 1 \end{array}\right].$$

Replace row2 by 3row1 + row2 and replace row3 by −2row1 + row3. This gives

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 8 & 10 & 3 & 1 & 0 \\ 0 & -6 & -5 & -2 & 0 & 1 \end{array}\right].$$

Replace row3 by $\frac{6}{8}$row2 + row3. This gives

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 8 & 10 & 3 & 1 & 0 \\ 0 & 0 & \frac{5}{2} & \frac{1}{4} & \frac{3}{4} & 1 \end{array}\right].$$

So, we have

$$U = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 8 & 10 \\ 0 & 0 & \frac{5}{2} \end{bmatrix}, L = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ \frac{1}{4} & \frac{3}{4} & 1 \end{bmatrix}^{-1} \quad \text{therefore } L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -\frac{3}{4} & 1 \end{bmatrix}.$$

This method for finding L should look familiar. However, don't mistake it for the procedure we used for finding the inverse of a matrix. When finding L, we performed Gaussian elimination. Then we found the inverse of the product of the elementary matrices. This inverse turned out to be L. This seems like more work than is necessary because we had to perform Gaussian elimination and then had to find the inverse of the matrix. There is a simpler method. Take a close look at L and at the scalar row multipliers used in the Gaussian elimination. When we eliminated the $a_{2,1}$ entry, we used the scalar 3. Notice that we have a $-3$ in the $l_{2,1}$ position of L. When we eliminated the $a_{3,1}$ entry, we used the scalar $-2$. Notice that we have a 2 in the $l_{3,1}$ position of L. When we eliminated the $a_{3,2}$ entry, we used the scalar $\frac{3}{4}$. Notice that we have a $-\frac{3}{4}$ in the $l_{3,2}$ position of L. These are not coincidences. To find the entries of L, we take the negative of the scalar row multipliers.

In the following example we will use the simpler method.

**Example IV.30** Given $A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ -3 & 2 & 1 & 0 \\ 2 & -2 & 1 & 0 \\ 1 & 2 & 3 & 4 \end{bmatrix}$, compute LU such that A = LU.

**Solution:** Write A and I. These will become L and U, respectively:

$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ -3 & 2 & 1 & 0 \\ 2 & -2 & 1 & 0 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Perform Gaussian elimination only on A and insert the negative of the scalar row multipliers into I.

Eliminate $a_{2,1}$: replace row2 by 3row1 + row2 and insert $-3$ into $l_{2,1}$.

Eliminate $a_{3,1}$: replace row3 by $-2$row1 + row3 and insert 2 into $l_{3,1}$.

Eliminate $a_{4,1}$: replace row4 by $-1$row1 + row4 and insert 1 into $l_{4,1}$.

After introducing 0's below the pivot in column 1, we have

$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 8 & 10 & 3 \\ 0 & -6 & -5 & -2 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Eliminate $a_{3,2}$: replace row3 by $\frac{6}{8}$row2 + row3 and insert $-\frac{6}{8}$ into $l_{3,2}$; since $a_{4,2}$ and $a_{4,3}$ require no elimination, insert zeros into $l_{4,2}$ and $l_{4,3}$, obtaining

$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 8 & 10 & 3 \\ 0 & 0 & \frac{5}{2} & \frac{1}{4} \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & -\frac{6}{8} & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \mathrm{UL}.$$

Gaussian elimination is now complete. Notice that, in the example above, we did not perform Gaussian elimination on the identity matrix. Instead, we inserted the negative of the scalar row multiplier into the identity matrix. If no scalar row multiplier was required to eliminate an entry in A, then the corresponding entry in the identity matrix was left as zero. It is crucial to remember that we do not perform row exchanges.

## G. BLOCK LU

The previous LU decomposition method used Gaussian elimination to factor the matrix A into the triangular matrices L and U, and works well when solving small systems of linear equations analytically. But what happens when our systems become too large to feasibly solve by hand? We must find a method that allows computers to do the work for us. By factoring A into upper and unit lower triangular matrices, we can develop a simple way to solve for $\mathbf{x}$ in the equation $\mathrm{A}\mathbf{x} = \mathbf{b}$. To see how to construct the LU factorization, consider the following block partitioning of the matrix equation $\mathrm{A} = \mathrm{LU}$,

$$A = \begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{a} & \widehat{A} \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{a}\alpha^{-1} & \widehat{L} \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{0} & \widehat{U} \end{bmatrix},$$

where $\widehat{L}$ and $\widehat{U}$ are lower and upper triangular matrices with smaller dimensions than the original A. Provided that $\widehat{L}\widehat{U} = \widehat{A} - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T$, we have the factorization we are looking for. The equation above is called a *Gauss step,* and shows how to get the first row and column of the LU factorization matrices. Once this is complete, we only need to compute the LU factorization of the *Schur complement* $\widehat{A} - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T$. This matrix is one row and one column smaller than the matrix we started with. Since it is clear that the LU factorization of a $1 \times 1$ matrix $A = [\alpha]$ is $[\alpha] = [1][\alpha]$, we can build a sequential procedure for finding the factorization. It is much easier to show how this algorithm works by showing a simple example. In the next example, to avoid confusion, we will use subscripts on A, L, and U.

**Example IV.31** Given $A = \begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix}$, compute LU such that LU = A.

**Solution**: Start by factoring A to find the first row and column in $L_0$ and $U_0$, so that

$$A = L_0U_0 = \begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{a} & A_1 \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{a}\alpha^{-1} & L_1 \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{0} & U_1 \end{bmatrix}.$$

Therefore,

$$\begin{aligned} \alpha &= 2, \mathbf{a} = \begin{bmatrix} -3 \\ 4 \end{bmatrix}, \mathbf{b}^T = \begin{bmatrix} 6 & 2 \end{bmatrix}, A_1 = \begin{bmatrix} -8 & 0 \\ 9 & 2 \end{bmatrix}, \text{ and} \\ A &= L_0U_0 = \begin{bmatrix} \underline{1} & \underline{0} & \underline{0} \\ -\underline{\tfrac{3}{2}} & & \\ \underline{2} & & L_1 \end{bmatrix} \begin{bmatrix} \underline{2} & \underline{6} & \underline{2} \\ \underline{0} & & \\ \underline{0} & & U_1 \end{bmatrix}. \end{aligned}$$

Compute the Schur complement $A_2 = A_1 - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T$, which gives

$$\begin{aligned} A_2 &= A_1 - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T = \begin{bmatrix} -8 & 0 \\ 9 & 2 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -3 \\ 4 \end{bmatrix}\begin{bmatrix} 6 & 2 \end{bmatrix} \\ &= \begin{bmatrix} -8 & 0 \\ 9 & 2 \end{bmatrix} - \begin{bmatrix} -9 & -3 \\ 12 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ -3 & -2 \end{bmatrix}. \end{aligned}$$

Factor $A_2$ to find the second row and column in $L_1$ and $U_1$, so that

$$\begin{aligned} A_2 &= \begin{bmatrix} 1 & 3 \\ -3 & -2 \end{bmatrix}, \alpha = 1, \mathbf{a} = [-3], \mathbf{b}^T = [3], \text{ and } A_3 = [-2], \text{ so} \\ A_2 &= \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{a}\alpha^{-1} & L_1 \end{bmatrix}\begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{0} & U_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{2} & \underline{1} & \underline{0} \\ 2 & -\underline{3} & L_2 \end{bmatrix}\begin{bmatrix} 2 & 6 & 2 \\ 0 & \underline{1} & \underline{3} \\ 0 & \underline{0} & U_2 \end{bmatrix}. \end{aligned}$$

Compute the Schur complement $A_4 = A_3 - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T$ so that

$$A_4 = A_3 - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T = [-2] - [-3][3] = [7].$$

Factor $A_4$ to find the third row and column in $L_2$ and $U_2$, which gives

$$\begin{aligned} A_4 &= [7], \alpha = 7, \mathbf{a} = [0], \mathbf{b}^T = [0], \text{ and } A_5 = [0], \text{ so} \\ A &= \begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{a} & \widehat{A} \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{a}\alpha^{-1} & \widehat{L} \end{bmatrix}\begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{0} & \widehat{U} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{2} & \underline{1} & 0 \\ 2 & -3 & \underline{1} \end{bmatrix}\begin{bmatrix} 2 & 6 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & \underline{7} \end{bmatrix} = LU. \end{aligned}$$

It turns out, however, that for a general matrix $A \in \mathcal{C}^{nxn}$, it is not always possible to find an LU factorization. If you want to prove this to yourself, try factoring

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

Proceeding in the normal manner, we immediately see that a divide-by-zero problem occurs when everything below the first pivot is divided by the pivot. Fortunately, it is always possible to find a permutation matrix P such that PA has an LU factorization. To begin, find a permutation $P_1$, and partition the matrix

$$P_1 A = \begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{a} & B \end{bmatrix}.$$

Choose $P_1$ such that either $\alpha \neq 0$, or both $\alpha = 0$ and $\mathbf{a} = 0$. Since the general structure of L and U is known, we partition them as

$$L = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{v} & \hat{L} \end{bmatrix} \text{ and } U = \begin{bmatrix} v & \mathbf{u}^T \\ \mathbf{0} & \hat{U} \end{bmatrix}.$$

Setting $P_1 A = LU$ gives

$$\begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{a} & B \end{bmatrix} = \begin{bmatrix} v & \mathbf{u}^T \\ v\mathbf{v} & \hat{L}\hat{U} + \mathbf{v}\mathbf{u}^T \end{bmatrix}.$$

Equating elements and simplifying yields

$$\begin{aligned} v &= \alpha \\ \mathbf{v} &= \frac{1}{\alpha}\mathbf{a} \\ \mathbf{u}^T &= \mathbf{b}^T \\ \hat{L}\hat{U} &= B - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T. \end{aligned}$$

The vector $\mathbf{u}^T$ can be taken directly from the permuted A, and both $\hat{L}$ and $\hat{U}$ come from factoring the Schur complement,

$$B - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T.$$

Of course, it can also happen that we must apply a permutation to the Schur complement to get a non-zero pivot. That is,

$$\hat{L}\hat{U} = \hat{P}_2(B - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T),$$

or

$$\widehat{\mathrm{P}}_2^T \widehat{\mathrm{L}} \widehat{\mathrm{U}} = \mathrm{B} - \frac{1}{\alpha}\mathbf{a}\mathbf{b}^T,$$

in which case

$$\mathrm{P}_1 \mathrm{A} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \frac{1}{\alpha}\mathbf{a} & \widehat{\mathrm{P}}_2^T \widehat{\mathrm{L}} \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{0} & \widehat{\mathrm{U}} \end{bmatrix}.$$

But this violates the structure of L, since $\widehat{\mathrm{P}}_2^T \widehat{\mathrm{L}}$ is not unit lower triangular. However, we notice that if we let

$$\mathrm{P}_2 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \widehat{\mathrm{P}}_2 \end{bmatrix},$$

then

$$\mathrm{P}_2 \mathrm{P}_1 \mathrm{A} = \mathrm{P}_2 \begin{bmatrix} 1 & \mathbf{0}^T \\ \frac{1}{\alpha}\mathbf{a} & \widehat{\mathrm{P}}_2^T \widehat{\mathrm{L}} \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{0} & \widehat{\mathrm{U}} \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \frac{1}{\alpha}\widehat{\mathrm{P}}_2 \mathbf{a} & \widehat{\mathrm{L}} \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{b}^T \\ \mathbf{0} & \widehat{\mathrm{U}} \end{bmatrix}.$$

This is a valid L and U. The reader can verify that the process can be continued until we have a complete LU factorization.

In practice, it is usually considered best to apply a permutation at every stage in the process and to choose it so that the pivot is as large as possible. In effect, we choose the permutation that puts the largest element of the first column into the first row. This strategy is called partial pivoting. Another strategy, full pivoting, utilizes both row and column interchanges to put the largest element remaining in the matrix into the first row and column.

## H. PIVOTING

We see in our derivation of the LU factorization that pivoting may be necessary to prevent division by zero. However, the importance of pivoting in a computer algorithm goes far beyond avoiding the divide by zero errors. As mentioned previously, pivoting is designed to get the largest number in a row or column into the pivot

position. Not only is this done to ensure that the pivot is not identically zero, but to ensure that the pivot is not very small. Extremely small pivots pose problems for two reasons.

The first reason is direct. If the pivot is very small, meaning much less (in magnitude) than 1, and the remaining numbers in the column are large, division by the pivot will produce very large numbers. These numbers pose a potential overflow problem in the computer. By employing partial pivoting to choose the largest element in the column as the pivot, this catastrophe is eliminated and guarantees that all remaining numbers in the column are less than one. Thus we can prevent an overflow error from occurring.

The second reason for pivoting is subtle, but far more important. If the pivot is small, one might wonder how it got that way. When performing the LU decomposition, if the first pivot of A is small, it just simply occupies the first pivot position. However, the remaining pivots are *computed* to be small due to the repeated subtraction in the Schur complement. Let's see what this really means. Say for example that we have two numbers, $a = \frac{1}{3}$ and $b = \frac{33301}{100000}$, and that our computer is a rounding machine that uses a base 10, 4 digit, floating point system. The computer will approximate these two numbers as $\hat{a} = .3333$ and $\hat{b} = .3330$. The *rounding errors* produced by these approximations look innocent enough since the approximations are so close to their actual values. But what happens when the machine computes $\hat{a} - \hat{b}$? Well, the machine represents this value as $.3000 \times 10^{-3}$, which is quite a bit different from the value $.3233 \times 10^{-3}$, the value closest to the true result without initial rounding. The phenomenon just shown, known as *catastrophic cancellation*, occurs when two very small numbers that are nearly equal are subtracted. The result leads to a great loss in precision. One way to lessen the effect of catastrophic cancellation is to avoid subtraction. That is why developers of computer algorithms spend a lot of time finding ways around the use of subtraction. However, subtraction is not the underlying problem, it is rounding that causes the problem of cancellation. One reasonable way

to avoid cancellation is to shy away from small numbers that were calculated from large numbers by successive subtractions. Such numbers are usually filled with error. Therefore, a small pivot is doubly dangerous. First, because it is probably filled with error, and second, because dividing by it magnifies the error. That is why it is so important to pivot when developing computer algorithms. This applies to solving systems of linear equations using both LU decomposition and Gaussian elimination algorithms.

## I. EXERCISES

1. Determine whether the following equations are linear or nonlinear in the variables $x, y$ and $z$. If nonlinear, state why.

a) $x + y + 2z = sin\pi$  b) $xy + z + 2 = 0$  c) $x^2 + y + \sqrt{5}z = 1$

d) $x = y$  e) $6x - 3y^{-1} = z$  f) $tan(\pi)x + e^2 y = z$  g) $sin(x) + y - 5z = 0$

2. Given the following systems of linear equations, use the method of substitution to compute the solution set.

a) $$\begin{array}{rcrcl} x_1 & + & x_2 & = & 1 \\ -2x_1 & + & 3x_2 & = & 4 \end{array}$$

b) $$\begin{array}{rcrcl} 3x_1 & - & 3x_2 & = & 2 \\ 6x_1 & + & x_2 & = & 1 \end{array}$$

c) $$\begin{array}{rcrcrcl} 2x_1 & - & 3x_2 & - & x_3 & = & 6 \\ x_1 & + & 6x_2 & - & 2x_3 & = & 12 \\ -x_1 & + & 4x_2 & + & 6x_3 & = & 24 \end{array}$$

3. Given the following system of linear equations,

$$
\begin{array}{rcrcrcr}
x_1 & + & 2x_2 & + & x_3 & = & 3 \\
2x_1 & + & 3x_2 & + & 4x_3 & = & 4 \\
4x_1 & + & 5x_2 & + & 9x_3 & = & 5,
\end{array}
$$

a) Write the augmented matrix.

b) Use Gaussian elimination to compute the solution set.

c) Verify the solution by substituting it back into the system of linear equations.

4. Given the following system of linear equations

$$
\begin{array}{rcrcrcr}
 & - & 2x_2 & + & 3x_3 & = & 1 \\
3x_1 & + & 6x_2 & - & 3x_3 & = & -2 \\
6x_1 & + & 6x_2 & + & 3x_3 & = & 5,
\end{array}
$$

a) Write the augmented matrix.

b) Use Gaussian elimination to compute the solution set.

c) Verify the solution by substituting it back into the system of linear equations.

5. Given the following system of linear equations,

$$
\begin{array}{rcrcrcrcr}
x & - & y & + & 2z & - & w & = & -1 \\
2x & + & y & - & 2z & - & 2w & = & -2 \\
-x & + & 2y & - & 4z & + & w & = & 1 \\
3x & & & & & - & 3w & = & -3,
\end{array}
$$

a) Write the augmented matrix.

b) Use Gaussian elimination to compute the solution set.

c) Verify the solution by substituting it back into the system of linear equations.

6. Given the following system of linear equations

$$\begin{array}{rcrcrcrcr} & - & x & + & 2y & + & 3z & = & 0 \\ w & + & x & + & 4y & + & 4z & = & 7 \\ w & + & 3x & + & 7y & + & 9z & = & 4 \\ -w & - & 2x & - & 4y & - & 6z & = & 6 \end{array}$$

a) Write the augmented matrix.

b) Use Gaussian elimination to compute the solution set.

c) Verify the solution set by substituting it back into the system of linear equations.

7. Given the following system of linear equations

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & 2x_3 & = & 5 \\ 2x_1 & + & 5x_2 & + & 7x_3 & = & 19 \\ 2x_1 & + & 4x_2 & + & 6x_3 & = & 16 \end{array}$$

a) Write the augmented matrix.

b) Use Gauss-Jordan elimination to compute the solution set.

c) Verify the solution set by substituting it back into the system of linear equations.

8. Given the following system of linear equations

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & x_3 & = & 6 \\ x_1 & + & 2x_2 & + & 2x_3 & = & 7 \\ 2x_1 & + & 4x_2 & + & 2x_3 & = & 15 \end{array}$$

a) Write the augmented matrix.

b) Use Gauss-Jordan elimination to compute the solution set.

c) Verify the solution set by substituting it back into the system of linear equations.

9. Given the following system of linear equations

$$\begin{array}{rcrcrcl} -x_1 & + & 4x_2 & + & x_3 & = & 3 \\ x_1 & + & 9x_2 & - & 2x_3 & = & 4 \\ 6x_1 & + & 4x_2 & - & 8x_3 & = & 5 \end{array}$$

a) Write the augmented matrix.

b) Use Gauss-Jordan elimination to compute the solution set.

c) Verify the solution set by substituting it back into the system of linear equations.

10. Given the following system of linear equations

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & x_3 & = & b_1 \\ 2x_1 & + & 3x_2 & + & 2x_3 & = & b_2 \\ 4x_1 & + & 5x_2 & + & 8x_3 & = & b_3 \end{array}$$

a) Write the augmented matrix.

b) Use Gaussian elimination to compute the solution set.

c) Can you find a vector $\mathbf{b}$, so that the augmented matrix is inconsistent?

d) Can you find a non-trivial solution if $\mathbf{b} = 0$?

11. Given the following system of linear equations

$$\begin{aligned} x_1 - 2x_2 - x_3 &= b_1 \\ -4x_1 + 5x_2 + 2x_3 &= b_2 \\ 4x_1 + 7x_2 + 6x_3 &= b_3 \end{aligned}$$

a) Write the augmented matrix.

b) Use Gaussian elimination to compute the solution set.

c) Can you find a vector **b**, so that the augmented matrix is inconsistent?

d) Can you find a non-trivial solution if $\mathbf{b} = 0$?

12. Given the following system of linear equations

$$\begin{aligned} u + 0 - w + 2x + y + z &= -3 \\ -3u + v + 0 - 6x - y - 4z &= 11 \\ 2u + 5v - 2w + 4x + 22y - 3z &= 4 \\ -u + 2v + 0 + 8x + 27y + 27z &= -33 \end{aligned}$$

a) Determine whether the system is consistent or inconsistent.

b) If the system is consistent, compute the solution set by any method.

13. Given $A = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$,

a) Compute $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

14. Given $A = \begin{bmatrix} 3 & 6 \\ 1 & 2 \end{bmatrix}$,

a) Compute $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

15. Given $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$,

a) Compute $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

16. Given $A = \begin{bmatrix} 1 & 2 & -1 \\ 4 & 7 & -3 \\ 2 & 1 & 3 \end{bmatrix}$,

a) Compute $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

17. Given $A = \begin{bmatrix} 1 & 6 & 2 \\ -2 & 3 & 5 \\ 7 & 12 & -4 \end{bmatrix}$,

a) Calculate $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

18. Given $A = \begin{bmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{bmatrix}$,

a) Compute $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

19. Given $A = \begin{bmatrix} 1 & 0 & 2 & 3 \\ -1 & 1 & 0 & 4 \\ 2 & 1 & -1 & 3 \\ -1 & 0 & 5 & 7 \end{bmatrix}$,

a) Compute $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

20. Given $A = \begin{bmatrix} i & 2 \\ 1 & -i \end{bmatrix}$,

a) Compute $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

21. Given $A = \begin{bmatrix} 1-i & 0 \\ 0 & 1+i \end{bmatrix}$,

a) Compute $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

22. Given $A = \begin{bmatrix} 1 & i & 0 \\ -i & 0 & 1 \\ 0 & 1+i & 1-i \end{bmatrix}$,

a) Compute $A^{-1}$, if it exists.

b) Is A singular, or nonsingular?

23. Given $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, compute $A^{-1}$.

24. Given $A = \begin{bmatrix} 4 & -2 & 7 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{bmatrix}$, compute $A^{-1}$.

25. Given $A = \begin{bmatrix} 2 & 0 & 0 \\ -6 & 0 & 0 \\ 4 & 3 & 6 \end{bmatrix}$, show that $A^{-1}$ does not exist.

26. Given upper and lower triangular matrices, show that each is invertible if, and only if, its diagonal elements are nonzero.

27. Given upper triangular matrices A and B, show that AB is also upper triangular.

28. Given $a_{1,1}, a_{1,2}, a_{2,1}$, and $a_{2,2} \in \mathcal{R}$, such that $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$, compute $b_{1,1}, b_{1,2}, b_{2,1}$, and $b_{2,2}$ such that $\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

29. Given $A = \begin{bmatrix} 3 & -6 & -3 \\ 2 & 0 & 6 \\ -4 & 7 & 4 \end{bmatrix}$, compute L and U such that A = LU.

30. Given $A = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 8 & 9 & 6 & 3 \\ 4 & 9 & 8 & 4 \\ 8 & 9 & 10 & 6 \end{bmatrix}$, compute L and U such that A = LU.

31. Given $L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -\frac{3}{4} & 1 \end{bmatrix}$, $U = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 8 & 10 \\ 0 & 0 & \frac{5}{2} \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, compute the solution set to $LU\mathbf{x} = \mathbf{b}$.

32. Given the system of linear equations $\begin{array}{rcrcrcr} 5x_1 & + & 5x_2 & + & 10x_3 & = & 0 \\ -8x_1 & - & 7x_2 & - & 9x_3 & = & 1 \\ & & 4x_2 & + & 26x_3 & = & 4 \end{array}$, use LU decomposition to compute the solution set.

33. Given the system of linear equations $\begin{array}{rcrcrcr} -x_1 & - & 3x_2 & - & 4x_3 & = & -6 \\ 3x_1 & + & 10x_2 & - & 10x_3 & = & -3 \\ -2x_1 & - & 4x_2 & + & 11x_3 & = & 9 \end{array}$, use LU decomposition to compute the solution set.

34. Given the system of linear equations
$$\begin{array}{rcrcrcl} 4x_1 & - & 2x_2 & + & x_3 & = & 2 \\ -12x_1 & + & 3x_2 & & & = & 5 \\ 8x_1 & - & x_2 & - & 2x_3 & = & 6 \end{array}$$
, use the Block Partitioning method of LU decomposition to compute the solution set.

# V. SOLUTIONS TO EXERCISES

## A. SOLUTIONS TO CHAPTER I

1a) $A^T = \begin{bmatrix} 3 & -6 & 1 \end{bmatrix}$ 1b) $B^H = \begin{bmatrix} 1-3i & -6 & 1+i \\ 2 & 0 & 5i \end{bmatrix}$

1c) $C^H = \begin{bmatrix} 2+6i & -6-4i & 1 \\ 7+4i & -5-6i & 3+i \\ 2 & -6-2i & 0 \end{bmatrix}$

2a) $A^T = \begin{bmatrix} 5+i & 7 & -2i \end{bmatrix}$ 2b) $B^H = \begin{bmatrix} 1+i & 3 \\ -4 & -7i \end{bmatrix}$

2c) $C^H = \begin{bmatrix} 2 & -6 & 14 \\ -4i & -5 & i \\ 22 & -3-4i & 10 \end{bmatrix}$

3a) $A + B = \begin{bmatrix} 8 \\ -7 \\ 9 \end{bmatrix}$ 3b) B and D are not compatibly-sized for addition.

3c) $D - C = \begin{bmatrix} 3+2i & -3+4i \\ -8-2i & 2-3i \\ 1+4i & 4-6i \end{bmatrix}$ 3d) $C^H$ and D are not compatibly-sized for addition.

3e) $2A - 3B = \begin{bmatrix} -9 \\ -9 \\ -22 \end{bmatrix}$ 3f) $iD^H + 2C^T = \begin{bmatrix} 8+3i & -4-12i & 5+3i \\ 6-8i & -3+2i & -6+9i \end{bmatrix}$

4a) $A - B = \begin{bmatrix} 5-i \\ 3 \\ -4 \end{bmatrix}$ 4b) B and C are not compatibly-sized for addition.

4c) $D^H + C = \begin{bmatrix} 4-3i & -5-4i \\ -2-4i & -14i \\ 1+i & 5 \end{bmatrix}$ 4d) $C^H - D = \begin{bmatrix} 4+i & 2 & -1-i \\ 11+4i & -14i & 5 \end{bmatrix}$

.

4e) $A + 2B = \begin{bmatrix} 5+2i \\ 21 \\ 26 \end{bmatrix}$ 4f) $D^H$ and $C^H$ are not compatibly-sized for subtraction.

6) $AB = \begin{bmatrix} 3 & 6 & -2 \\ 6 & 12 & -5 \\ 9 & 18 & -8 \end{bmatrix} \neq \begin{bmatrix} 6 & 6 & 6 \\ -19 & -20 & -21 \\ 15 & 18 & 21 \end{bmatrix} = BA$ Yes, in general $AB \neq BA$.

7a) $BA = \begin{bmatrix} -1 & 0 & 0 \\ 4 & 8 & -2 \\ 1 & -4 & -4 \end{bmatrix}$ 7b) $B^TC = \begin{bmatrix} 7-18i & 24-13i & 6 \\ -9+2i & 7i & -10+4i \\ -26+16i & -26+26i & -24+8i \end{bmatrix}$

7c) $C^H B = \begin{bmatrix} 7+18i & -9-2i & -26-16i \\ 24+13i & -7i & -26-26i \\ 6 & -10-4i & -24-8i \end{bmatrix}$ 7d) $a(BA) = \begin{bmatrix} -\alpha & 0 & 0 \\ 4\alpha & 8\alpha & -2\alpha \\ \alpha & -4\alpha & -4\alpha \end{bmatrix}$

7e) $BA + BC = \begin{bmatrix} -1-14i & 16-6i & 2i \\ -4+8i & 10+8i & -14+4i \\ -5-2i & -8+4i & -8+2i \end{bmatrix}$

8a) $AB = \begin{bmatrix} 15 & 33 & 41 \\ 9 & 24 & 19 \end{bmatrix}$ 8b) $BA^T = \begin{bmatrix} 32 & 28 \\ 16 & 14 \\ 24 & 15 \end{bmatrix}$

8c) $C^H B = \begin{bmatrix} 3+5i & 9+16i & 5+7i \\ 7-2i & 14-4i & 21-6i \end{bmatrix}$ 8d) $a(BC) = \begin{bmatrix} a(8-24i) & a(16+2i) \\ a(4-12i) & a(8+i) \\ -13ai & a(17+i) \end{bmatrix}$

8e) AB and C are not compatibly-sized for addition.

## B. SOLUTIONS TO CHAPTER II

1a) $\mathbf{x}^H = \begin{bmatrix} -1-2i & 2+3i & 7i \end{bmatrix}$ 1b) $\mathbf{y}^T = \begin{bmatrix} -2 & 3 & 1 \end{bmatrix}$ 1c) $\mathbf{x}+\mathbf{y} = \begin{bmatrix} -3+2i \\ 5-3i \\ 1-7i \end{bmatrix}$

1d) $3\mathbf{x}-2i\mathbf{y} = \begin{bmatrix} -3+10i \\ 6-15i \\ -23i \end{bmatrix}$ 1e) $\mathbf{x}$ and $\mathbf{z}$ are not compatibly-sized for addition.

1f) $\mathbf{x}+\mathbf{z}^T = \begin{bmatrix} -1+i \\ -7i \\ 6-7i \end{bmatrix}$ 1g) $2\mathbf{z}^H = \begin{bmatrix} -2i \\ -4-8i \\ 12 \end{bmatrix}$ 1h) $\mathbf{z}+\overline{\mathbf{z}} = \begin{bmatrix} 0 & -4 & 12 \end{bmatrix}$

3a) $\mathbf{x}\cdot\mathbf{z} = 1$ 3b) $\mathbf{x}\cdot\mathbf{y} = 5$

3c) $\mathbf{x}(\mathbf{z}\cdot\mathbf{y}) = \begin{bmatrix} -19 \\ -38 \\ -76 \end{bmatrix}$ 3d) $(\mathbf{y}\cdot\mathbf{z})\mathbf{x} = \begin{bmatrix} -19 \\ -38 \\ -76 \end{bmatrix}$

4a) $\mathbf{x}\cdot\mathbf{z} = 39$ 4b) $(\mathbf{x}\cdot\mathbf{z})\mathbf{y} = \begin{bmatrix} -78 \\ 117 \\ 39 \end{bmatrix}$

4c) $\mathbf{x}(\mathbf{z}\cdot\mathbf{y}) = \begin{bmatrix} -155 \\ 62 \\ 31 \end{bmatrix}$ 4d) $(\mathbf{y}\cdot\mathbf{z})\mathbf{x} = \begin{bmatrix} -155 \\ 62 \\ 31 \end{bmatrix}$

6a) $32+8i$ 6b) $17-17i$ 6c) $38+5i$ 6d) $32-8i$ 6e) $17+17i$
6f) $38-5i$

7a) $\begin{bmatrix} -78 \\ 117 \\ 39 \end{bmatrix}$ 7b) $\begin{bmatrix} -155 \\ 62 \\ 31 \end{bmatrix}$ 7c) $\begin{bmatrix} -85 \\ 119 \\ 0 \end{bmatrix}$ 7d) $\begin{bmatrix} -78 \\ 117 \\ 39 \end{bmatrix}$

8a) $\begin{bmatrix} 76+10i \\ 109-53i \\ 53+109i \end{bmatrix}$ 8b) $\begin{bmatrix} -102-68i \\ 85-17i \\ -51+85i \end{bmatrix}$ 8c) $\begin{bmatrix} -176-24i \\ -32-128i \\ 96-24i \end{bmatrix}$ 8d) $\begin{bmatrix} -76+10i \\ 109-53i \\ 53+109i \end{bmatrix}$

9) $\mathbf{x}\cdot\mathbf{y} = 10$, not orthogonal $\mathbf{x}\cdot\mathbf{z} = 0$, orthogonal $\mathbf{x}\cdot\mathbf{w} = 0$, orthogonal
$\mathbf{y}\cdot\mathbf{z} = 2$, not orthogonal $\mathbf{y}\cdot\mathbf{w} = 0$, orthogonal $\mathbf{z}\cdot\mathbf{w} = -14$, not orthogonal

10a) $\|\mathbf{x}\| = \sqrt{(4)^2 + (-2)^2 + (2)^2} = \sqrt{24}$ 10b) $\| \mathbf{y} \| = \sqrt{62}$ 10c) $\|\mathbf{z}\| = \sqrt{30}$
10d) $\mathbf{x}$ and $\mathbf{y}$ are not orthogonal. 10e) $\mathbf{z}$ and $\mathbf{y}$ are not orthogonal.
10f) $\mathbf{x}$ and $\mathbf{z}$ are orthogonal.

11a) $\|\mathbf{x}\| = \sqrt{x^H x} = \sqrt{59}$ 11b) $\|\mathbf{y}\| = \sqrt{116}$ 11c) $\|\mathbf{z}\| = \sqrt{23}$
11d) $\mathbf{x}$ and $\mathbf{y}$ are not orthogonal. 11e) $\mathbf{z}$ and $\mathbf{y}$ are orthogonal.
11f) $\mathbf{x}$ and $\mathbf{z}$ are not orthogonal.

12a) $\sqrt{94}$ 12b) $\sqrt{296}$ 12c) $\sqrt{80}$ 12d) $\sqrt{90}$ 12e) $\mathbf{w}$ and $\mathbf{x}$, $\mathbf{x}$ and $\mathbf{y}$

13a) $\mathbf{w}^T\mathbf{v} = -16$ 13b) $\mathbf{w}^T\mathbf{w} = 54$ 13c) $\cos\theta = \frac{\mathbf{w}^H\mathbf{v}}{\|\mathbf{w}\|\|\mathbf{v}\|} = -\frac{16}{\sqrt{14}\sqrt{54}} = -\frac{8\sqrt{21}}{63}$

13d) $\text{proj}_{\mathbf{w}}\mathbf{v} = \mathbf{w}\frac{\mathbf{w}^T\mathbf{v}}{\mathbf{w}^T\mathbf{w}} = \begin{bmatrix} 3 \\ -5 \\ -4 \\ 2 \end{bmatrix}\left(-\frac{16}{54}\right)$

14a) $\mathbf{w}^T\mathbf{v} = -33$ 14b) $\mathbf{w}^T\mathbf{w} = 50$ 14c) $\cos\theta = \frac{-33}{\sqrt{26}\sqrt{50}}$

14d) $\text{proj}_{\mathbf{w}}\mathbf{v} = \begin{bmatrix} 3 \\ -5 \\ -4 \\ 0 \end{bmatrix}\left(-\frac{33}{50}\right)$

15a) $\frac{-14+13i}{\sqrt{51}\sqrt{43}}$ 15b) $\frac{18+2i}{\sqrt{51}\sqrt{14}}$ 15c) $\frac{-16}{\sqrt{14}\sqrt{54}}$ 15d) $\frac{-31-6i}{\sqrt{43}\sqrt{54}}$ 15e) $\frac{7+5i}{\sqrt{14}\sqrt{43}}$
15f) $\frac{25-28i}{\sqrt{51}\sqrt{54}}$

16) $\mathbf{x} = \frac{3}{4}\mathbf{u} - \frac{3}{4}\mathbf{w} - \frac{1}{2}\mathbf{v}$

17a) $\mathbf{v} = c_1\mathbf{w} + c_2\mathbf{x} + c_3\mathbf{y}$ 17b) $\mathbf{w} + \mathbf{x} + \mathbf{y} = \mathbf{v}$ which is equivalent to setting all of the constants in part a equal to 1. Therefore, $\mathbf{v}$ is in the span $\{\mathbf{w}, \mathbf{x}, \mathbf{y}\}$
17c) No, they are linearly dependent because $\mathbf{v}$ depends on a combination of the other 3 vectors.

18) No, $\mathbf{b}$ is not a linear combination of the columns of A because the linear system is inconsistent.

19) Yes, $\mathbf{b}$ is a linear combination of the columns of A because the linear system is consistent.

## C. SOLUTIONS TO CHAPTER III

1. Check $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ and $T(c\mathbf{x}) = cT(\mathbf{x})$.

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) &= T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \end{bmatrix}\right) = \begin{bmatrix} (x_1+y_1)-(x_2+y_2) \\ 0 \\ (x_1+y_1+x_2+y_2)^2 \end{bmatrix} \\ &= \begin{bmatrix} x_1+y_1-x_2-y_2 \\ 0 \\ x_1^2+y_1^2+x_2^2+y_2^2+2x_1y_1+2x_1x_2+2x_2y_2+2x_1y_2+2y_1y_2+2x_2y_1 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} T(\mathbf{x})+T(\mathbf{y}) &= T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) + T\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1-x_2 \\ 0 \\ (x_1+x_2)^2 \end{bmatrix} + \begin{bmatrix} y_1-y_2 \\ 0 \\ (y_1+y_2)^2 \end{bmatrix} \\ &= \begin{bmatrix} x_1+y_1-x_2-y_2 \\ 0 \\ x_1^2+y_1^2+x_2^2+y_2^2+2x_1x_2+2y_1y_2 \end{bmatrix} \end{aligned}$$

Therefore, $T(\mathbf{x}+\mathbf{y}) \neq T(\mathbf{x}) + T(\mathbf{y})$.

$$T(c\mathbf{x}) = T\left(c\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} cx_1 \\ cx_2 \end{bmatrix}\right) = \begin{bmatrix} cx_1-cx_2 \\ (cx_1+cx_2)^2 \end{bmatrix} = c\begin{bmatrix} x_1-x_2 \\ c(x_1+x_2)^2 \end{bmatrix}$$

$$cT(\mathbf{x}) = cT\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = c\begin{bmatrix} x_1 - x_2 \\ (x_1 + x_2)^2 \end{bmatrix}$$

$T(c\mathbf{x}) \neq cT(\mathbf{x})$. Neither rule holds true, therefore, $T$ is not a linear transformation.

2) Check $T(\mathbf{x}+\mathbf{y})$ $=T(\mathbf{x})+T(\mathbf{y})$ and $T(c\mathbf{x}) = cT(\mathbf{x})$.

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) &= T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}\right) = \begin{bmatrix} 3(x_1 + y_1) \\ 2(x_1 + y_1) - (x_2 + y_2) \\ x_2 + y_2 \end{bmatrix} \\ &= \begin{bmatrix} 3x_1 + 3y_1 \\ 2x_1 + 2y_1 - x_2 - y_2 \\ x_2 + y_2 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} T(\mathbf{x}) + T(\mathbf{y}) &= T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) + T\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 \\ 2x_1 - x_2 \\ x_2 \end{bmatrix} + \begin{bmatrix} 3y_1 \\ 2y_1 - y_2 \\ y_2 \end{bmatrix} \\ &= \begin{bmatrix} 3x_1 + 3y_1 \\ 2x_1 + 2y_1 - x_2 - y_2 \\ x_2 + y_2 \end{bmatrix} \end{aligned}$$

Therefore, $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$.

$$T(c\mathbf{x}) = T\left(c\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} cx_1 \\ cx_2 \end{bmatrix}\right) = \begin{bmatrix} 3cx_1 \\ 2cx_1 - cx_2 \\ cx_2 \end{bmatrix} = c\begin{bmatrix} 3x_1 \\ 2x_1 - x_2 \\ x_2 \end{bmatrix}$$

$$cT(x) = cT\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = c\begin{bmatrix} 3x_1 \\ 2x_1 - x_2 \\ x_2 \end{bmatrix}$$

$T(c\mathbf{x}) = cT(\mathbf{x})$. Both rules hold true, therefore, $T$ is a linear transformation.

3) Check $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ and $T(c\mathbf{x}) = cT(\mathbf{x})$.

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) &= T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1+y_1 \\ x_1+y_1+4(x_2+y_2) \\ (x_1+y_1)(x_2+y_2) \end{bmatrix} \\ &= \begin{bmatrix} x_1+y_1 \\ x_1+y_1+4x_2+4y_2 \\ x_1x_2+y_1x_2+x_1y_2+y_1y_2 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} T(\mathbf{x})+T(\mathbf{y}) &= T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) + T\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1+4x_2 \\ x_1x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_1+4y_2 \\ y_1y_2 \end{bmatrix} \\ &= \begin{bmatrix} x_1+y_1 \\ x_1+y_1+4x_2+4y_2 \\ x_1x_2+y_1y_2 \end{bmatrix} \end{aligned}$$

Therefore, $T(\mathbf{x}+\mathbf{y}) \neq T(\mathbf{x}) + T(\mathbf{y})$.

$$T(c\mathbf{x}) = T\left(c\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} cx_1 \\ cx_2 \end{bmatrix}\right) = \begin{bmatrix} cx_1 \\ cx_1+4cx_2 \\ c^2x_1x_2 \end{bmatrix} = c\begin{bmatrix} x_1 \\ x_1+4x_2 \\ cx_1x_2 \end{bmatrix}$$

$$cT(\mathbf{x}) = cT\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = c\begin{bmatrix} x_1 \\ x_1+4x_2 \\ x_1x_2 \end{bmatrix}$$

$T(c\mathbf{x}) \neq cT(\mathbf{x})$. Neither rule holds true, therefore, $T$ is not a linear transformation.

4a) $\|A\|_1 = 14$  4b) $\|B\|_1 = 14$  4c) $\| C \|_1 = 13$

5a) $\|A\|_\infty = 13$  5b) $\|B\|_\infty = 18$  5c) $\| C \|_\infty = 10$

6a) elementary matrix (multiplies row2 by $-4$)

6b) not an elementary matrix

6c) elementary matrix (interchanges row2 and row3)

6d) elementary matrix (replaces row1 by $-3$row3 + row1)

6e) not an elementary matrix

6f) elementary matrix (multiplies row1 by $-2$)

7a) $E_1$ interchanges rows 1 and 3 of a compatibly-sized matrix. $E_2$ replaces row1 of a compatibly-sized matrix by (row2 − row1). $E_3$ multiplies row2 of a compatibly-sized matrix by $-2$.

7b) $E_1E_2E_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ -1 & -2 & 0 \end{bmatrix}$  7c) $E_3E_2E_1 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

7d) No, because matrix multiplication is generally not commutative.

8a) Independent  8b) Dependent  8c) 4  8d) 4  8e) 4

9a) $\text{rank}(A) = 2$  9b) $\mathbf{x} = [4r + 28s + 37t - 13u, 2r + 12s + 16t - 5u, r, s, t, u]^T$

# D. SOLUTIONS TO CHAPTER IV

1a) linear  1b) nonlinear (xy)  1c) nonlinear ($x^2$)

1d) linear  1e) nonlinear ($3y^{-1}$)  1f) linear  1g) nonlinear (sin(x))

2a) $\mathbf{x} = \begin{bmatrix} -\frac{1}{5} & \frac{6}{5} \end{bmatrix}^T$  2b) $\mathbf{x} = \begin{bmatrix} \frac{5}{21} & -\frac{3}{7} \end{bmatrix}^T$  2c) $\mathbf{x} = \begin{bmatrix} 8 & 2 & 4 \end{bmatrix}^T$

3a) $\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 2 & 3 & 4 & 4 \\ 4 & 5 & 9 & 5 \end{array}\right]$  3b) $\mathbf{x} = \begin{bmatrix} -6 \\ 4 \\ 1 \end{bmatrix}$

4a) $\left[\begin{array}{ccc|c} 0 & -2 & 3 & 1 \\ 3 & 6 & -3 & -2 \\ 0 & -6 & 9 & 9 \end{array}\right]$  4b) No solution. System is inconsistent.

5a) $\left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 2 & 1 & -2 & -2 & -2 \\ -1 & 2 & -4 & 1 & 1 \\ 3 & 0 & 0 & -3 & -3 \end{array}\right]$  5b) $\begin{bmatrix} t-1 \\ 2s \\ s \\ t \end{bmatrix}$, where $s, t \in \mathcal{R}$

6a) $\left[\begin{array}{rrrr|r} 0 & -1 & 2 & 3 & 0 \\ 1 & 1 & 4 & 4 & 7 \\ 1 & 3 & 7 & 9 & 4 \\ -1 & -2 & -4 & -6 & 6 \end{array}\right]$ 6b) $\begin{bmatrix} -\frac{82}{13} \\ 1 \\ \frac{128}{13} \\ -\frac{85}{13} \end{bmatrix}$

7a) $\left[\begin{array}{rrr|r} 1 & 1 & 2 & 5 \\ 2 & 5 & 7 & 19 \\ 2 & 4 & 6 & 16 \end{array}\right]$ 7b) $\begin{bmatrix} 2-s \\ 3-s \\ s \end{bmatrix}$, where $s \in \mathcal{R}$.

8a) $\left[\begin{array}{rrr|r} 1 & 2 & 1 & 6 \\ 1 & 2 & 2 & 7 \\ 2 & 4 & 2 & 15 \end{array}\right]$ 8b) No solution. System is inconsistent.

9a) $\left[\begin{array}{rrr|r} -1 & 4 & 1 & 3 \\ 1 & 9 & -2 & 4 \\ 6 & 4 & -8 & 5 \end{array}\right]$ 9b) $\begin{bmatrix} \frac{133}{2} \\ \frac{9}{2} \\ \frac{103}{2} \end{bmatrix}$

10a) $\left[\begin{array}{rrr|r} 1 & 2 & 1 & b_1 \\ 2 & 3 & 2 & b_2 \\ 4 & 5 & 8 & b_3 \end{array}\right]$ 10b) $\begin{bmatrix} \frac{26b_1-9b_2+b_3}{4} \\ 2b_1-b_2 \\ \frac{-6b_1+b_2+b_3}{4} \end{bmatrix}$ 10c) No 10d) No

11a) $\left[\begin{array}{rrr|r} 1 & -2 & -1 & b_1 \\ -4 & 5 & 2 & b_2 \\ 4 & 7 & 8 & b_3 \end{array}\right]$ 11b) $\begin{bmatrix} \frac{-5b_1-2b_2-3t}{3} \\ \frac{-4b_1-b_2-2t}{3} \\ t \end{bmatrix}$, where $t \in \mathcal{R}$

11c) $15b_1 + 5b_2 + b_3 \neq 0$ 11d) $b = \begin{bmatrix} -t \\ -\frac{2t}{3} \\ t \end{bmatrix}$, where $t \in \mathcal{R}$.

12a) The system has an infinite number of solutions and is therefore, consistent.

12b) $\begin{bmatrix} \frac{-110-58s+50t}{10} & 2+-29s+t & -9s & \frac{40-21s-30t}{10} & s & t \end{bmatrix}^T$, where $s, t \in \mathcal{R}$.

13a) $\text{A}^{-1} = \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix}$ 13b) Nonsingular

14a) $\text{A}^{-1} = \begin{bmatrix} 3 & 6 \\ 1 & 2 \end{bmatrix}$ 14b) Singular

15a) $\text{A}^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 2 & -2 & -1 \\ -1 & 1 & 1 \end{bmatrix}$ 15b) Nonsingular

16a) $\text{A}^{-1} = \begin{bmatrix} -12 & \frac{7}{2} & -\frac{1}{2} \\ 9 & -\frac{5}{2} & \frac{1}{2} \\ -5 & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$ 16b) Nonsingular

17a) $A^{-1}$ does not exist. 17b) Singular

18a) $A^{-1} = \begin{bmatrix} 0 & 1 & 0 & 2 \\ 1 & -1 & -2 & 2 \\ 0 & 1 & 3 & -3 \\ -2 & 2 & 3 & -2 \end{bmatrix}$ 18b) Nonsingular

19a) $A^{-1}$ does not exist. 19b) Singular

20a) $A^{-1} = \begin{bmatrix} i & 2 \\ 1 & -i \end{bmatrix}$ 20b) Nonsingular

21a) $A^{-1} = \begin{bmatrix} \frac{1+i}{2} & 0 \\ 0 & \frac{1-i}{2} \end{bmatrix}$ 21b) Nonsingular

22a) $A^{-1} = \begin{bmatrix} \frac{1+i}{2} & \frac{1+i}{2} & \frac{i}{2} \\ \frac{-1-i}{2} & \frac{-1+i}{2} & \frac{-i}{2} \\ \frac{-1+i}{2} & \frac{1+i}{2} & \frac{-1}{2} \end{bmatrix}$ 22b) Nonsingular

23) $A^{-1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

24) $A^{-1} = \begin{bmatrix} \frac{1}{4} & \frac{2}{20} & \frac{33}{20} \\ 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & -1 \end{bmatrix}$

25) Gaussian elimination will produce a row of zeros.

26) Gaussian elimination will produce a row of zeros.

28) $a_{1,1}b_{1,1} + a_{1,2}b_{2,1} = 1$ and $a_{2,1}b_{1,1} + a_{2,2}b_{2,1} = 0 \longrightarrow b_{2,1} = -\frac{a_{2,1}b_{1,1}}{a_{2,2}}$.
Substituting $b_{2,1}$ into the first equation yields

$$a_{1,1}b_{1,1} + a_{1,2}\frac{-a_{2,1}b_{1,1}}{a_{2,2}} = 1 \longrightarrow b_{1,1} = \frac{a_{2,2}}{a_{1,1}b_{2,2} - a_{2,1}a_{1,2}}.$$

Substituting $b_{1,1}$ into $b_{2,1} = -\frac{a_{2,1}b_{1,1}}{a_{2,2}}$ yields

$$b_{2,1} = -\frac{a_{2,1}}{a_{1,1}b_{2,2} - a_{2,1}a_{1,2}}.$$

$a_{2,1}b_{1,2} + a_{2,2}b_{2,2} = 1$ and $a_{1,1}b_{1,2} + a_{1,2}b_{2,2} = 0 \longrightarrow b_{2,2} = -\frac{a_{1,1}b_{1,2}}{a_{1,2}}$.
Substituting $b_{2,2}$ into the first equation yields

$$a_{2,1}b_{1,2} + a_{2,2}\frac{-a_{1,1}b_{1,2}}{a_{1,2}} = 1 \longrightarrow b_{1,2} = \frac{-a_{2,1}}{a_{1,1}b_{2,2} - a_{2,1}a_{1,2}}.$$

Substituting $b_{1,2}$ into $b_{2,2} = \frac{-a_{1,1}b_{1,2}}{a_{1,2}}$ yields

$$b_{1,2} = \frac{-a_{1,2}}{a_{1,1}b_{2,2} - a_{2,1}a_{1,2}}.$$

29) $L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 1 & 0 \\ -\frac{4}{3} & -\frac{1}{4} & 1 \end{bmatrix}$ and $U = \begin{bmatrix} 3 & -6 & -3 \\ 0 & 4 & 8 \\ 0 & 0 & 2 \end{bmatrix}$

30) $L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 2 & 1 & 2 & 1 \end{bmatrix}$ and $U = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

31) $\mathbf{z} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} -\frac{2}{5} \\ -\frac{1}{2} \\ \frac{4}{5} \end{bmatrix}$

32) $L = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{8}{5} & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}$, $U = \begin{bmatrix} 5 & 5 & 10 \\ 0 & 1 & 7 \\ 0 & 0 & -2 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

33) $L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix}$, $U = \begin{bmatrix} -1 & -3 & -4 \\ 0 & 1 & -22 \\ 0 & 0 & 63 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$.

34) $L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix}$, $U = \begin{bmatrix} 4 & -2 & 1 \\ 0 & -3 & 3 \\ 0 & 0 & -1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -\frac{55}{12} \\ -\frac{50}{3} \\ -13 \end{bmatrix}$